

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:41:54 ; Search time 3.2668 Seconds
(Without alignments)
7582.941 Million cell updates/sec

Title: US-09-902-772-1_COPY_650_660
Perfect score: 11
Sequence: 1 agagagagagag 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
25: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	11	100.0	21	22	AAH49300
C 2	11	100.0	21	24	ABK94127
C 3	11	100.0	21	24	ABK94128
C 4	11	100.0	22	18	AA75373
C 5	11	100.0	22	19	AAV59955
C 6	11	100.0	23	14	AAQ39034
C 7	11	100.0	24	24	AAI81861
C 8	11	100.0	25	22	AAH37422
C 9	11	100.0	31	14	AAQ40991

C 10	11	100.0	32	18	AA791953	PCR primer 1 used
C 11	11	100.0	36	22	AA788029	Human Trp 1 PCR pr
C 12	11	100.0	40	21	AA295968	Polynucleotide seq
C 13	11	100.0	43	20	AAK81695	Primer used in a n
C 14	11	100.0	51	21	AAH7247	Human clone c94397
C 15	11	100.0	60	24	ABN40086	Human spliced tran
C 16	11	100.0	63	19	AAV29642	Nucleotide sequenc
C 17	11	100.0	63	22	AA500974	PCR primer for con
C 18	11	100.0	65	24	ABN58012	Mouse spliced tran
C 19	11	100.0	72	13	AAQ33600	Microsatellite seq
C 20	11	100.0	76	13	AAQ34136	Sequence upstream
C 21	11	100.0	90	24	ABK76766	Bacillus lichenifo
C 22	11	100.0	98	24	AA563133	Cell death protect
C 23	11	100.0	99	22	ABA48277	Human breast cell
C 24	11	100.0	99	22	ABA66159	Human foetal liver
C 25	11	100.0	99	22	ABA33228	Probe #11694 for g
C 26	11	100.0	99	22	AAK14578	Human brain expres
C 27	11	100.0	99	22	AAK40324	Human bone marrow
C 28	11	100.0	99	22	AAK12082	Probe #11015 for g
C 29	11	100.0	99	22	AAI46343	Probe #15029 used
C 30	11	100.0	99	22	AAI06802	Probe #6793 used t
C 31	11	100.0	99	24	AB514344	Human genome-deriv
C 32	11	100.0	111	14	AAQ39030	Oligonucleotide id
C 33	11	100.0	111	14	AAQ39031	Oligonucleotide id
C 34	11	100.0	111	14	ABK40333	Probe #18799 for g
C 35	11	100.0	113	21	AAV61549	Microsatellite mar
C 36	11	100.0	113	19	AACT7053	Human secreted pro
C 37	11	100.0	115	14	AAQ39032	Oligonucleotide id
C 38	11	100.0	115	14	AB514344	Oligonucleotide id
C 39	11	100.0	117	24	AB520896	Human genome-deriv
C 40	11	100.0	118	19	AAV16144	Microsatellite mar
C 41	11	100.0	124	22	ABA47414	Human breast cell
C 42	11	100.0	124	22	ABA65302	Human foetal liver
C 43	11	100.0	124	22	ABA32403	Probe #10869 for g
C 44	11	100.0	124	22	AAK13716	Human brain expres
C 45	11	100.0	124	22	AAK39457	Human bone marrow

ALIGNMENTS

RESULT 1	
AAH49300/c	
ID	AAH49300 standard; DNA; 21 BP.
XX	
XX	AAH49300;
XX	
XX	28-NOV-2001 (first entry)
XX	
DE	V. vinifera aquaporin probe SEQ ID 12.
XX	
KW	Aquaporin; transmembrane protein; intrinsic protein; water transport;
KW	cell growth; drought tolerance; salt tolerance; probe; ss.
XX	
OS	Vitis vinifera.
XX	
PN	W0200166793-A2.
XX	
PD	13-SEP-2001.
XX	
PF	09-MAR-2001; 2001WO-EP02718.
XX	
PR	09-MAR-2000; 2000DE-1011480.
XX	
PA	(GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEIT.
XX	
PI	Baiges Blanco IM, Schaeffner AR, Mas A;
XX	
XX	WPI; 2001-589950/66.
XX	
DR	
XX	
PT	Nucleic acid for amplifying and quantitatively and qualitatively
PT	detecting aquaporin genes; gene fragments; RNA or expression of an
PT	aquaporin gene from Vitis species

XX Claim 1; Page 28; 38pp; German.
XX
CC This invention describes a novel nucleic acid for amplifying and
CC quantitatively and qualitatively detecting aquaporin genes, RNA or gene
CC expression. The nucleic acid is used to qualitatively and/or
CC quantitatively detect or amplify aquaporin genes, gene fragments or gene
CC expression from Vitis species. Aquaporin is a transmembrane protein
CC involved in water transport and known to affect cell growth and drought
CC and salt tolerance. This sequence represents a probe used in the
CC detection of the Vitis vinifera aquaporin associated with the intrinsic
CC protein family.
XX
SQ Sequence 21 BP; 2 A; 9 C; 1 G; 9 T; 0 other;
XX
Query Match 100.0%; Score 11; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAGAG 11
Db 20 AGAGAGAGAGAG 10
RESULT 2
ABK94127 ID ABK94127 standard; DNA; 21 BP.
XX
AC ABK94127;
XX
DT 27-AUG-2002 (first entry)
XX
DE Endothelin-2 (EDN-2) SNP detection PCR primer #23.
XX
KW Endothelin: EDN; endothelin converting enzyme; ECE; endothelin receptor;
KW EDNR; signaling system; cardiovascular disease; coronary heart disease;
KW hypertension; atherosclerosis; angiogenesis; fatty acid metabolism;
KW diabetes; familial hypercholesterolemia; forensic marker;
KW transgenic animal; solid support; cardiovascular regulator; SNP;
KW single nucleotide polymorphism; PCR; primer; ss.
XX
OS Synthetic.
XX
PN WO200224747-A2.
XX
PD 28-MAR-2002.
XX
PE 31-AUG-2001; 2001WO-EP10087.
XX
PR 19-SEP-2000; 2000EP-0120123.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Brinkmann U, Hoffmeyer S;
XX
DR WPI; 2002-435060/46.
XX
PT Novel polynucleotide of the endothelin/endothelin converting
PT enzyme/receptors of endothelin and endothelin converting enzyme
PT signaling system associated with cardiovascular disease, useful for
PT treating the disease
XX
Claim 1; Page 57; 190pp; English.
XX
XX The invention describes a polynucleotide (I) of the endothelin
CC (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR)
CC signaling system which is associated with a cardiovascular disease. (I),
CC the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I)
CC or (II) is useful for producing cells capable of expressing a molecular
CC variant polypeptide which is associated with a cardiovascular disease.
CC (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing
CC a molecular variant gene comprising (I) is useful for identifying and
CC obtaining a pro-drug or drug capable of modulating the activity of a

CC molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system
CC or its gene product, or for identifying and obtaining an inhibitor of
CC the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE
CC signaling system or its gene product. The isolated proteins and
CC polynucleotides encoding them are useful for preparation of a
CC pharmaceutical composition for treating a cardiovascular disease such as
CC coronary heart disease, hypertension, atherosclerosis, or related to
CC abnormal angiogenesis or fatty acid metabolism e.g. diabetes and familial
CC hypercholesterolemia. The gene or a polynucleotide fragment of the
CC EDN/ECE/EDNR signaling system are useful as forensic markers, for
CC creating a transgenic animal and in creation of a solid support
CC comprising polynucleotides, genes, vectors, polypeptides, antibodies or
CC host cells of the invention. This sequence represents a PCR primer used
CC to identify single nucleotide polymorphisms in DNA encoding
CC cardiovascular regulator proteins of the EDN/ECE/EDNR signaling pathway.
XX
SQ Sequence 21 BP; 7 A; 4 C; 10 G; 0 U; 0 other;
XX
Query Match 100.0%; Score 11; DB 24; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAGAG 11
Db 11 AGAGAGAGAGAG 21
RESULT 3
ABK94128/C ID ABK94128 standard; DNA; 21 BP.
XX
AC ABK94128;
XX
DT 27-AUG-2002 (first entry)
XX
DE Endothelin-2 (EDN-2) SNP detection PCR primer #24.
XX
KW Endothelin: EDN; endothelin converting enzyme; ECE; endothelin receptor;
KW EDNR; signaling system; cardiovascular disease; coronary heart disease;
KW hypertension; atherosclerosis; angiogenesis; fatty acid metabolism;
KW diabetes; familial hypercholesterolemia; forensic marker;
KW transgenic animal; solid support; cardiovascular regulator; SNP;
KW single nucleotide polymorphism; PCR; primer; ss.
XX
OS Synthetic.
XX
PN WO200224747-A2.
XX
PD 28-MAR-2002.
XX
PE 31-AUG-2001; 2001WO-EP10087.
XX
PR 19-SEP-2000; 2000EP-0120123.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Brinkmann U, Hoffmeyer S;
XX
DR WPI; 2002-435060/46.
XX
PT Novel polynucleotide of the endothelin/endothelin converting
PT enzyme/receptors of endothelin and endothelin converting enzyme
PT signaling system associated with cardiovascular disease, useful for
PT treating the disease
XX
Claim 1; Page 57; 190pp; English.
XX
XX The invention describes a polynucleotide (I) of the endothelin
CC (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR)
CC signaling system which is associated with a cardiovascular disease. (I),
CC the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I)
CC or (II) is useful for producing cells capable of expressing a molecular
CC variant polypeptide which is associated with a cardiovascular disease.

CC (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing
CC a molecular variant gene comprising (I) is useful for identifying and
CC obtaining a pro-drug or drug capable of modulating the activity of a
CC molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system
CC or its gene product, or for identifying and obtaining an inhibitor of
CC the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE
CC signaling system or its gene product. The isolated proteins and
CC polynucleotides encoding them are useful for preparation of a
CC pharmaceutical composition for treating a cardiovascular disease such as
CC coronary heart disease, hypertension, atherosclerosis, or related to
CC abnormal angiogenesis or fatty acid metabolism e.g. diabetes and familial
CC hypercholesterolaemia. The gene or a polynucleotide fragment of the
CC EDN/ECE/EDNR signaling system are useful as forensic markers, for
CC creating a transgenic animal and in creation of a solid support
CC comprising polynucleotides, genes, vectors, polypeptides, antibodies or
CC host cells of the invention. This sequence represents a PCR primer used
CC to identify single nucleotide polymorphisms in DNA encoding
CC cardiovascular regulator proteins of the EDN/ECE/EDNR signaling pathway.
XX
SQ Sequence 21 BP; 0 A; 10 C; 4 G; 7 T; 0 other;

Query Match 100.0%; Score 11; DB 24; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11
11 AGAGAGAGAG 1

RESULT 4
AAAT75373 standard; CDNA; 22 BP.

AAAT75373;

24-DEC-1998 (first entry)

CDNA synthesis primer EGRI-6.

ss; human; RAD50; DNA repair; tumour suppression; cancer; Septin-2;
central nervous system; PCR; primer; amplification.

Synthetic.

MO9727284-A2.

31-JUL-1997.

24-JAN-1997; 97WO-US01299.

17-JUL-1996; 96US-0687080.

26-JAN-1996; 96US-0592126.

(GENE-) GENELABS TECHNOLOGIES INC.

Dolganov G;

WPI; 1997-393672/36.

Human tumour suppressor gene RAD50 - useful to detect
predisposition to, decrease risk of and treat cancer, also Septin-2
homologues

Example 1; Page 36; 155pp; English.

The primers AAAT75354-75378 were used to for cDNA synthesis in the
method of the invention. Disclosed in the invention is human RAD50
(RAD50) which is involved in DNA repair and has tumour suppression
activity and can be used to detect predisposition to, decrease the risk
of or treat cancers, e.g. acute myeloid leukaemia, myelodysplastic
syndrome, therapy related myelodysplastic syndrome, therapy related
acute myeloid leukaemia, refractory anaemia or refractory anaemia with

CC excess blasts. Also disclosed in this invention are human Septin-2
CC homologues which may be used as targets for cancer therapies and central
CC nervous system directed treatment methods, and to measure the
CC proliferative potential of selected cell types.

SQ Sequence 22 BP; 2 A; 12 C; 0 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 18; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11
15 AGAGAGAGAG 5

RESULT 5
AAV59955 standard; DNA; 22 BP.

AAV59955;

25-NOV-1998 (first entry)

PCR primer EGRI-6 used to amplify EGRI-1 cDNA.

Human analogue; yeast RAD50; Drosophila Septin-2; Acyl-CoA synthetase;
immunomodulatory activity; identification; activated T-cell; cytokine;
EGRI-1; PCR primer; ss.

Synthetic.

Homo sapiens.

MO9838306-A1.

03-SEP-1998.

27-FEB-1997; 97WO-US03159.

27-FEB-1997; 97WO-US03159.

(GENE-) GENELABS TECHNOLOGIES INC.

Dolganov G;

WPI; 1998-481207/41.

Novel human immunomodulatory poly(peptide)s - have homology to the
yeast RAD50 or Drosophila Septin-2 proteins

Example 1; Page 27; 155pp; English.

PCR primers AAV59955-56 were used to identify cDNA encoding human
cytokine EGRI-1 from different cDNA pools, to provide an estimate of
the degree to which the cytokine transcript is present. mRNA was
isolated from activated T-cells, and converted to cDNA prior to
amplification. The specification describes sequences encoding human
analogues of the yeast RAD50, the Drosophila Septin-2 and Acyl-CoA
synthetase. The proteins have immunomodulatory activity. The nucleic
acids and proteins can be used to identify activated T-cells in a sample
population. They can also be used to isolate and identify sequences
encoding other proteins or other compounds having immunomodulatory
activity.

Sequence 22 BP; 2 A; 12 C; 0 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11
15 AGAGAGAGAG 5

```

RESULT 6
AAQ39034
ID AAQ39034 standard; CDNA; 23 BP.
XX
AC AAQ39034;
XX
DT 22-JUL-1993 (first entry)
XX
DE Oligonucleotide identifier tag 3' primer.
XX
KW Oligonucleotide; identifier; tag; primer; oligomer; library; agonist;
KW identification; ligand; antagonist; cell membrane receptor; hormone;
KW toxin; venom; viral epitope; sugar; cofactor; peptide; drug; protein;
KW enzyme substrate; PCR; amplify; ss.
XX
OS Synthetic.
XX
PN WO9306121-A.
XX
PD 01-APR-1993.
XX
PF 16-SEP-1992; 92WO-US07815.
XX
PR 18-SEP-1991; 91US-0762522.
XX
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Barrett RW, Dower WJ, Gallop MA;
XX
DR WPI; 1993-117464/14.
XX
PT Synthetic oligomer library to facilitate identification of
PT oligomers with required properties - comprises several members,
PT each of which has an oligomer of monomer sequence linked to
PT identifier tag
XX
PS Disclosure; Page 37; 76pp; English.
XX
CC This sequence is a primer which was used in the preparation of the
CC sequences given in AAQ39030-33. The amplified oligonucleotides are
CC identifier tags. They were prepared using this primer and also a
CC primer corresponding to bases 7-28 of the sequence given in AAQ39030.
CC The tags were then sequenced using the primer sequences given in
CC AAQ39035-36. These identifier tags were used in the construction of
CC an oligomer library comprising >10,000,000 members. The use of the
CC identification tags on the oligomer members of the library
CC facilitates identification of oligomers with required properties.
CC Ligands which can be investigated include agonists and antagonists
CC for cell membrane receptors, toxins, venoms, viral epitopes, hormones,
CC sugars, cofactors, peptides, enzyme substrates, drugs and proteins.
XX
SQ Sequence 23 BP; 11 A; 0 C; 12 G; 0 U; 0 other;
Query Match 100.0%; Score 11; DB 14; Length 23;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGGAG 11
DB 10 AGAGAGAGGAG 20

```

```

XX
KW Cardiovascular disease; TagIB polymorphism; coronary heart disease; ss;
KW cholesteryl ester transfer protein; CERP; Tag I restriction site;
KW B1 allele; myocardial infarction; angina pectoris; coronary death;
KW coronary artery disease; coronary insufficiency; PCR primer; human.
XX
OS Homo sapiens.
XX
PN WO200185999-A1.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US15013.
XX
PR 11-MAY-2000; 2000US-203467P.
XX
PA (TUFT) TUFTS COLLEGE.
XX
PI Ordovas JM, Schaefer EJ;
XX
DR WPI; 2002-082902/11.
XX
PF Assessing risk for the development of cardiovascular disease in an
PT individual, comprises analysing nucleic acid from the individual for
PT presence of TagIB polymorphism of cholesteryl ester transfer protein
PT gene.
XX
PS Claim 13; Page 67; 72pp; English.
XX
CC The invention relates to a method for assessing risk for development of
CC cardiovascular disease, comprising analysing a nucleic acid from an
CC individual for the presence of TagIB polymorphism of cholesteryl ester
CC transfer protein (CERP) gene, and assessing the risk for development of
CC cardiovascular disease. The method involves determining whether the
CC individual is homozygous for, heterozygous for or does not possess the
CC TagIB polymorphism, whereby the absence of the polymorphism correlates
CC with the frequency and phenotypic expression of coronary heart disease.
CC The invention also includes PCR primers used for amplification of a
CC suitable section of the first intron of the CERP gene encompassing the
CC TagI restriction site of the B1 allele of the CERP gene, the presence of
CC the TagI restriction site being indicative of the absence of the TagIB
CC polymorphism. The cardiovascular diseases include myocardial infarction,
CC coronary artery disease, angina pectoris, coronary insufficiency and
CC coronary death. This sequence represents a PCR primer used in the method
CC of the invention.
XX
SQ Sequence 24 BP; 7 A; 7 C; 8 G; 2 T; 0 other;
Query Match 100.0%; Score 11; DB 24; Length 24;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGGAG 11
DB 10 AGAGAGAGGAG 20

```

```

RESULT 8
AAH37422/C
ID AAH37422 standard; DNA; 25 BP.
XX
AC AAH37422;
XX
DT 14-AUG-2001 (first entry)
XX
DE SNP specific lower PCR primer SEQ ID 218.
XX
KW Single nucleotide polymorphism; SNP; single nucleotide primer extension;
KW SNPE; genotyping; agammaglobulinaemia; diabetes insipidus; cancer;
KW Lesch-Nyhan syndrome; muscular dystrophy; familial hypercholesterolaemia;
KW polycystic kidney disease; osteogenesis imperfecta; autoimmune disease;
KW acute intermittent porphyria; rheumatoid arthritis; multiple sclerosis;
KW inflammation; forensic investigation; paternity analysis; PCR primer; ss.

```

XX OS Homo sapiens.
XX PN NO200129262-A2.
XX PD 26-APR-2001.
XX PF 13-OCT-2000; 2000MO-US28436.
XX PR 15-OCT-1999; 99US-0160096.
XX PA (ORCH-) ORCHID BIOSCIENCES INC.
XX PI Picoult-Newburg L, Pohl M;
XX WPI; 2001-290930/30.
XX DR
XX PT New genotyping oligonucleotide, useful for detecting the presence,
XX PT absence or identity of single polynucleotide polymorphism in a nucleic
XX PT acid sample
XX PS
XX PS Claim 1; Page 51; 83pp: English.
XX CC Sequences AAH37205 - AAH40944 represent PCR primers, single nucleotide
XX CC primer extension (SNPE) primers, and the sequences of regions flanking
XX CC sites of single nucleotide polymorphisms SNPs. The present invention
XX CC includes kits for determining the presence or absence of a SNP, using the
XX CC oligonucleotides of the invention. The PCR primers are used to amplify a
XX CC SNP flanking sequence, the SNPE primer is used as a genotyping primer.
XX CC The oligonucleotides are useful for genotyping a nucleic acid sample by
XX CC performing a single-nucleotide primer extension reaction. The
XX CC oligonucleotides are useful for determining the presence, absence or
XX CC identity of a SNP and for genotyping nucleic acid samples, for e.g. to
XX CC assess by association analysis the genotype of an individual or group of
XX CC individuals, having a pathological phenotypic trait suspected of being
XX CC caused by one or more SNPs. Phenotypic traits include diseases e.g.
XX CC agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular
XX CC dystrophy, familial hypercholesterolemia, polycystic kidney disease,
XX CC osteogenesis imperfecta and acute intermittent porphyria. Phenotypic
XX CC traits also include symptoms of or susceptibility to multifactorial
XX CC disease of which a component is or may be genetic such as autoimmune
XX CC diseases, including, rheumatoid arthritis, multiple sclerosis,
XX CC inflammation, cancer, nervous system diseases and infection by pathogenic
XX CC microorganism. The method is also useful in forensic investigations and
XX CC paternity analysis. The present sequence represents a PCR primer specific
XX CC for a human SNP containing DNA sequence.
XX SQ
XX SQ Sequence 25 BP; 4 A; 12 C; 0 G; 9 T; 0 other;
Query Match 100.0%; Score 11; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGGAG 11
Db 22 AGAGAGAGGAG 12
RESULT 9
AAQ40991
ID AAQ40991 standard; DNA; 31 BP.
XX AAQ40991;
XX AC
XX XX
XX DT 17-DEC-2001 (updated)
XX DT 06-OCT-1993 (first entry)
XX DE Mono ADP-ribosyltransferase gene isolation sequence.
XX DE
XX KM Antisense technology; gene therapy; structure/function studies;
XX KM production; preparation; prodn.; prepn.; antibody; NAD: arginine; ss.
XX OS Synthetic.

XX XX USN7985698-N.
XX PN 01-MAY-1993.
XX PD 30-NOV-1992; 92US-0985698.
XX PF 30-NOV-1992; 92US-0985698.
XX PR 30-NOV-1992; 92US-0985698.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX PI Moss J, Nightingale MS, Okazaki I, Zolkiewska A;
XX WPI; 1993-188869/23.
XX DR
XX XX Gene sequences encoding mono-ADP-ribosyl:transferase -
XX PT specifically from rabbit skeletal muscle and human form, useful
XX PT in antisense technology
XX PS
XX PS Example; Page 42; 46pp: English.
XX CC The sequence is given in the specification but is not
XX CC specifically defined.
XX CC (Note: Revised entry submitted to correct the patent number format of
XX CC US Government-owned NPLS applications to prevent clashes with ongoing US
XX CC granted patent numbers. For further information please visit the Derwent
XX CC web site at www.derwent.com/dwpi/updates/npls-us.html.)
XX SQ
XX SQ Sequence 31 BP; 12 A; 0 C; 19 G; 0 U; 0 other;
Query Match 100.0%; Score 11; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGGAG 11
Db 5 AGAGAGAGGAG 15
RESULT 10
AAT91953/c
ID AAT91953 standard; cDNA; 32 BP.
XX AAT91953;
XX AC
XX XX 20-MAR-1998 (first entry)
XX DT
XX DE PCR primer 1 used to localize the antigenic peptide recognised by TIL586.
XX DE
XX KM Tyrosinase related protein 1 gene; TRP-1; gp75; tumour antigen;
XX KM tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
XX KM alternative reading frame; cancer detection; pre-cancer detection;
XX KM melanoma; PCR primer; ss.
XX OS
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9729195-A2.
XX PD 14-AUG-1997.
XX PF 06-FEB-1997; 97MO-US02186.
XX PR 04-OCT-1996; 96US-0725736.
XX PR 09-FEB-1996; 96US-0595602.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Rosenberg SA, Wang R;
XX WPI; 1997-415349/38.
XX XX Cancer antigen peptide(s) derived from the tyrosinase-related

PT protein 1 or 2 - useful for detecting, preventing or treating a
 PT cancer in a mammal, especially melanoma
 XX
 PS Example 1; Page 34; 111pp; English.
 XX
 CC PCR primers AAT91953-54 were used to amplify the tyrosinase related
 CC protein 1 gene (TRP-1 or gp75, AAT91944), from positions 1-667 in order
 CC to localise the region containing the predominant T-cell epitope.
 CC The PCR products contain an internal ATG codon in frame with the TRP-1
 CC gene, located at 445 bp as a start codon for translation of a truncated
 CC normal TRP-1 protein. However, neither of the PCR products conferred the
 CC ability to stimulate cytokine secretion from TIL586, suggesting that the
 CC epitope is located upstream of these fragments (see AAT91949-51). Novel
 CC cancer peptides have been identified in TRP-1 and TRP-2. The nucleic
 CC acids encoding the cancer peptides or TRP-2 can be used to detect a
 CC cancer or pre-cancer in a mammal, especially by detecting the presence
 CC of the alternative ORF 3 of the TRP-1 gene or the sequence encoding the
 CC novel tumour antigen TRP-2. Vectors and recombinant viruses containing
 CC antigen peptide encoding nucleic acids, antibodies raised against the
 CC peptides, or the peptides themselves can be used to prevent or treat a
 CC cancer in a mammal, especially a melanoma.
 XX
 SQ Sequence 32 BP; 7 A; 9 C; 7 G; 9 T; 0 other;
 Query Match 100.0%; Score 11; DB 18; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGAGAGAGGAG 11
 |||||||||
 Db 29 AGAGAGAGGAG 19
 RESULT 11
 AAF88029/c
 ID AAF88029 standard; DNA: 36 BP.
 XX
 AC AAF88029;
 XX
 DT 16-JUL-2001 (first entry)
 XX
 DE Human Trp 1 PCR primer trp1-5/2-LIGA.
 XX
 KW Listeria: expression vector; tumor-associated antigen; Trp 1; Trp 2;
 KW MelanA/MART-1; cytostatic; attenuated; immunotherapy; malignant melanoma;
 KW pigmented tumor; malignant schwannoma; vaccination; tyrosinase;
 KW antigen-presenting cell; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200127295-A1.
 XX
 PD 19-APR-2001.
 XX
 PF 13-OCT-2000; 2000MO-DE03629.
 XX
 PR 14-OCT-1999; 99DE-1049594.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSGEZENTRUM.
 XX
 PI Schandendorf D, Paschen A, Chakraborty T, Domann E;
 DR WPI; 2001-282041/29.
 XX
 PT Listeria expression vector for immunotherapy, particularly of malignant
 PT melanoma, comprises a DNA sequence encoding tumor-associated antigens -
 PS Examples; Page 16; 41pp; German.
 XX
 CC This invention describes a novel Listeria expression vector (A) for
 CC immunotherapy which comprises a promoter (P) functional in Listeria,
 CC operably linked to a DNA sequence (I) encoding one of the
 CC tumor-associated antigens (II) human tyrosinase, Trp 1 or 2, or

CC MelanA/MART-1. The products of the invention have cytostatic activity.
 CC Recombinant attenuated Listeria containing (A) are useful for
 CC immunotherapy (prophylactic, adjuvant or therapeutic), specifically of
 CC malignant melanoma (but also other pigmented tumors such as malignant
 CC schwannoma), particularly as a replacement for radiotherapy. Using
 CC attenuated Listeria as carrier for (A) provides a simple way of
 CC vaccination, since antigen-presenting cells acquire tumor-associated
 CC antigens by natural infection, eliminating the need for labor-intensive
 CC ex vivo modification of autologous cells. This sequence represents the
 CC PCR primer trp1-5/2-LIGA used in the amplification of the human
 CC Trp 1 which is described in the method of the invention.
 XX
 SQ Sequence 36 BP; 8 A; 13 C; 5 G; 10 T; 0 other;
 Query Match 100.0%; Score 11; DB 22; Length 36;
 Best Local Similarity 100.0%; Pred. No. 9.3e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGAGAGAGGAG 11
 |||||||||
 Db 35 AGAGAGAGGAG 25
 RESULT 12
 AAZ95968
 ID AAZ95968 standard; DNA: 40 BP.
 XX
 AC AAZ95968;
 XX
 DT 10-APR-2000 (first entry)
 XX
 DE Polynucleotide sequence including binding site for BamHI.
 XX
 KW Ligand binding; restriction enzyme; nucleic acid determination;
 KW pharmaceutical; BamHI. ss.
 XX
 OS Synthetic.
 XX
 PN WO963077-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 04-JUN-1999; 99MO-US12516.
 XX
 PR 04-JUN-1998; 98US-0087905.
 XX
 PR 03-JUN-1999; 99US-0324672.
 XX
 PA (TMTE-) TM TECHNOLOGIES INC.
 XX
 PI Lane MJ, Benight AS, Faldasz BD;
 DR WPI; 2000-116369/10.
 XX
 PT Modulating polynucleotide ligand binding site affinity using
 PT determination of the flanking duplex sequences -
 PS Example 1; Page 40; 62pp; English.
 XX
 CC The invention provides a method for determining the sequence of
 CC polynucleotide flanking regions that modulate ligand binding
 CC characteristics of an adjacent binding site. The method comprises:
 CC (i) providing a number of different duplex polynucleotides, each having
 CC the same polynucleotide ligand binding site and a randomly synthesised
 CC sequence flanking the binding site; (ii) exposing the duplex to a ligand
 CC selective for the binding site; (iii) isolating duplexes which bind or
 CC do not bind the ligand, and (iv) determining the nucleotide composition
 CC of the flanking duplex sequence by sequencing the duplex sequence
 CC adjacent to the binding site. The invention is used to modulate the
 CC ligand-binding characteristics of any nucleotide sequence. The invention
 CC is less costly and more efficient than prior art techniques that moderate
 CC ligand binding using small molecule pharmaceuticals. Sequences
 CC AAZ95762-256170 represent polynucleotide sequences including the binding
 CC site for the restriction enzyme BamHI and used in the course of the

CC Invention.
XX
SQ Sequence 40 BP; 12 A; 7 C; 14 G; 7 T; 0 other;
Query Match 100.0%; Score 11; DB 21; Length 40;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAGAG 11
DB 28 AGAGAGAGAGAG 38
RESULT 13
AAx81695
ID AAx81695 standard; DNA; 43 BP.
AC AAx81695;
XX
DT 03-SEP-1999 (first entry)
XX
DE Primer used in a novel method of DNA sequencing.
XX
KW Extension region; terminal nucleotide; template arranging segment;
KW complexity decreasing nucleotide; DNA sequencing; primer; ss.
XX
OS Synthetic.
XX JP1151092-A.
XX
PD 08-JUN-1999.
XX
PF 24-AUG-1998; 98JP-0237840.
XX
PR 22-AUG-1997; 97US-0916120.
XX
PA (LYNX-) LYNX THERAPEUTICS INC.
XX WPI: 1999-388481/33.
XX
PT Extension of DNA using rolling primer - useful as DNA sequencing
PT method
XX
PS Disclosure; Page 26; 32pp; Japanese.
XX
CC The specification describes a method for the determination of the
CC nucleotide sequence of a polynucleotide. The method comprises providing
CC a set of primers in which each primer has an extension region containing
CC a terminal nucleotide, a template arranging segment and at least one
CC complexity-decreasing nucleotide, forming a template containing
CC primer-combining sites and the polynucleotide in which the
CC primer-combining sites are complementary to at least one primer of the
CC set, forming an amplicon from the template by amplifying a
CC double-stranded DNA formed selectively by extending the primer from
CC the set in which the extending region forms a double-strand completely
CC matched to primer-combining sites of the template, identifying the
CC terminal nucleotide of the extending region of the primer by an
CC identification of the amplicon, shifting the primer-combining sites by
CC one nucleotide to the direction of extension by varying the
CC primer-combining sites of the template, and repeating this until the
CC nucleotide sequence of the polynucleotide is determined. The method can
CC be used for DNA sequencing. The present sequence represents a primer
CC used to demonstrate the invention.
XX
SQ Sequence 43 BP; 15 A; 2 C; 17 G; 1 T; 8 other;
Query Match 100.0%; Score 11; DB 20; Length 43;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAGAG 11
DB 13 AGAGAGAGAGAG 23

RESULT 14
AAx77247
ID AAx77247 standard; cDNA; 51 BP.
XX
AC AAx77247;
XX
DT 16-NOV-2000 (first entry)
XX
DE Human clone c943979039 polymorphic site, SEQ ID NO:930.
XX
KW Human; single nucleotide polymorphism; SNP; chromosome 2;
KW detection; identification; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace (26,A)
FT /*tag=a
XX
XX WO200029623-A2.
XX
PD 25-MAY-2000.
XX
PF 17-NOV-1999; 99WO-US27293.
XX
PR 17-NOV-1998; 98US-0109024.
PR 16-NOV-1999; 99US-0109024.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinketsu RA, Leach MD;
XX
DR WPI: 2000-387826/33.
XX
PT Human nucleic acids containing single nucleotide polymorphisms, useful
PT for treating a subject suffering, or at risk from a pathology due to
PT the presence of a sequence polymorphism -
XX
PS Claim 1; Page 438; 543pp; English.
XX
CC Sequences AAx76318-A77509 represent 1192 human nucleic acid sequences
CC which contain single nucleotide polymorphisms (SNPs). Sequences 1 to
CC 1112 (AAx76318-A77429) are consecutive pairs of nucleotides which
CC contain silent SNPs. Sequences 1113 to 1192 (AAx77430-A77509) are
CC consecutive pairs of nucleotides containing SNPs which result in changes
CC in the corresponding amino acid sequences (AAx11749-R11828). The SNPs in
CC sequences 1113 to 1128 (AAx77430-A77445) lead to conservative amino acid
CC changes, while those in sequences 1129 to 1186 (AAx77446-A77503) result
CC in non-conservative changes. The SNPs in sequences 1187 to 1192
CC (AAx77504-A77509) generate frameshift mutations. The invention also
CC relates to a method of detecting a polymorphic site in a nucleic acid
CC a method of determining the relatedness of two nucleic acids. It also
CC encompasses peptides containing polymorphic sites, antibodies raised
CC against such peptides, and a method of detecting polymorphic
CC proteins/peptides using the antibodies. The nucleic acids are useful for
CC gene therapy of an individual having, suspected of having, or at risk of
CC developing a pathological condition due to the presence of a sequence
CC polymorphism. Such treatment would comprise administration of the
CC wild-type nucleic acid sequence. Antibodies raised against polymorphic
CC peptides can also be used in the treatment of such individuals.
XX
SQ Sequence 51 BP; 13 A; 20 C; 12 G; 6 T; 0 other;
Query Match 100.0%; Score 11; DB 21; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAGAG 11
DB 16 AGAGAGAGAGAG 26

Search completed: July 28, 2003, 11:28:03
 Job time : 6.2668 secs

```

RESULT 15
ID ABN40086/C
XX ABN40086 standard; DNA; 60 BP.
AC ABN40086;
XX
XX
XX 15-JUL-2002 (first entry)
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:12834.
DE
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-1B01903.
XX
XX 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX
XX WPI; 2002-257383/30.
XX
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX
XX Example 1; SEQ ID 12834; 47pp; English.
XX
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.
XX
XX
XX Sequence 60 BP; 5 A; 25 C; 8 G; 22 T; 0 other;
SQ
Query Match 100.0%; Score 11; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAGAG 11
| | | | | | | | | |
Db 45 AGAGAGAGAGAG 35

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:00:14 ; Search time 0.694102 Seconds
(without alignments)
4860.160 Million cell updates/sec

Title: US-09-902-772-1_COPY_650_660

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2.6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2.6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	11	100.0	22	1	US-08-592-126-24
C 2	11	100.0	22	2	US-08-687-080-24
C 3	11	100.0	23	1	US-07-946-239-5
C 4	11	100.0	23	3	US-09-151-467-5
C 5	11	100.0	23	3	US-09-036-599-10
C 6	11	100.0	23	4	US-09-256-838-5
C 7	11	100.0	23	5	PCR-US92-07815-5
C 8	11	100.0	30	4	US-09-147-915-6
C 9	11	100.0	32	2	US-08-599-602-42
C 10	11	100.0	32	3	US-09-197-816-42
C 11	11	100.0	36	2	US-08-863-639A-26
C 12	11	100.0	43	1	US-08-560-313A-8
C 13	11	100.0	43	1	US-08-611-155B-12
C 14	11	100.0	43	2	US-08-916-120A-14
C 15	11	100.0	63	4	US-08-951-822-12
C 16	11	100.0	63	4	US-09-173-043-26
C 17	11	100.0	63	4	US-09-368-951-12
C 18	11	100.0	90	1	US-08-314-536-34
C 19	11	100.0	90	3	US-08-320-982-34
C 20	11	100.0	90	3	US-08-819-037-34
C 21	11	100.0	90	4	US-09-045-940-34
C 22	11	100.0	104	1	US-08-314-536-34
C 23	11	100.0	104	1	US-08-320-982-31
C 24	11	100.0	104	3	US-08-819-037-31
C 25	11	100.0	104	4	US-09-045-940-31
C 26	11	100.0	111	1	US-08-488-470A-9
C 27	11	100.0	111	1	US-08-488-470A-10

C 28	11	100.0	111	1	US-07-946-239-1	Sequence 1, Appl1
C 29	11	100.0	111	1	US-07-946-239-2	Sequence 2, Appl1
C 30	11	100.0	111	1	US-08-484-505A-9	Sequence 9, Appl1
C 31	11	100.0	111	1	US-08-484-505A-10	Sequence 10, Appl1
C 32	11	100.0	111	3	US-09-151-467-1	Sequence 1, Appl1
C 33	11	100.0	111	3	US-09-151-467-2	Sequence 2, Appl1
C 34	11	100.0	111	3	US-09-036-599-5	Sequence 5, Appl1
C 35	11	100.0	111	3	US-09-036-599-6	Sequence 6, Appl1
C 36	11	100.0	111	4	US-09-078-403A-9	Sequence 9, Appl1
C 37	11	100.0	111	4	US-09-078-403A-10	Sequence 10, Appl1
C 38	11	100.0	111	4	US-09-256-838-1	Sequence 1, Appl1
C 39	11	100.0	111	4	US-09-256-838-2	Sequence 2, Appl1
C 40	11	100.0	111	5	PCR-US92-07815-1	Sequence 1, Appl1
C 41	11	100.0	111	5	PCR-US92-07815-2	Sequence 2, Appl1
C 42	11	100.0	115	1	US-08-488-470A-11	Sequence 11, Appl1
C 43	11	100.0	115	1	US-08-488-470A-12	Sequence 12, Appl1
C 44	11	100.0	115	1	US-07-946-239-3	Sequence 3, Appl1
C 45	11	100.0	115	1	US-07-946-239-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-592-126-24/c
Sequence 24, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO.: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Primer EGR1-6
US-08-592-126-24
Query Match 100.0%; Score 11; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11

Db 15 AGAGAGAGAG 5

RESULT 2

US-08-687-080-24/C
Sequence 24, Application US/08687080
Patent No. 5963427
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Primer EGR11-6
US-08-687-080-24

Query Match 100.0%; Score 11; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11
Db 15 AGAGAGAGAG 5

RESULT 3
US-07-946-239-5
Sequence 5, Application US/07946239
Patent No. 5770358
GENERAL INFORMATION:
APPLICANT: DOWER, WILLIAM J
APPLICANT: BARRETT, RONALD W
APPLICANT: GALLOP, MARK A
APPLICANT: NEEDELS, MICHAEL C
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE
COLLECTIONS OF OLIGOMERS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND AND TOWNSEND
STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,239
FILING DATE: 19920916
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11509-36-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-5043
TELEFAX: 415-543-9600
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-07-946-239-5

Query Match 100.0%; Score 11; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGAGAGAGAG 11
Db 10 AGAGAGAGAG 20

RESULT 4

US-09-151-467-5
Sequence 5, Application US/09151467
Patent No. 6140493
GENERAL INFORMATION:
APPLICANT: DOWER, WILLIAM J
APPLICANT: BARRETT, RONALD W
APPLICANT: GALLOP, MARK A
APPLICANT: NEEDELS, MICHAEL C
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE
COLLECTIONS OF OLIGOMERS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND
STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,467
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/946,239
FILING DATE: 1992-09-16
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.

REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11509-36-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-151-467-5

Query Match 100.0%; Score 11; DB 3; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

oy 1 AGAGAGAGGAG 11
|||||
Db 10 AGAGAGAGGAG 20

RESULT 5
US-09-036-599-10
Sequence 10, Application US/09036599
Patent No. 6143497
GENERAL INFORMATION:
APPLICANT: DOWER, WILLIAM J.
APPLICANT: BARRETT, RONALD W.
APPLICANT: GALLOP, MARK A.
TITLE OF INVENTION: Method of Synthesizing Diverse
TITLE OF INVENTION: Collections of Oligomers
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESS: William M. Smith
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,599
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/762,522
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11509-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-5043
TELEFAX: 415-326-2400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-036-599-10

Query Match 100.0%; Score 11; DB 3; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

oy 1 AGAGAGAGGAG 11
|||||
Db 10 AGAGAGAGGAG 20

RESULT 6
US-09-256-838-5
Sequence 5, Application US/09256838
Patent No. 6416949
GENERAL INFORMATION:
APPLICANT: DOWER, WILLIAM J.
APPLICANT: BARRETT, RONALD W.
APPLICANT: GALLOP, MARK A.
NEDELS, MICHAEL C
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE
TITLE OF INVENTION: COLLECTIONS OF OLIGOMERS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESS: TOWNSEND AND TOWNSEND
STREET: 1 MARKET PLAZA, STEWART TOWER, SUITE 2000
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/256,838
FILING DATE: 24-Feb-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/946,239
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11509-36-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-256-838-5

Query Match 100.0%; Score 11; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

oy 1 AGAGAGAGGAG 11
|||||
Db 10 AGAGAGAGGAG 20

RESULT 7
PCT-US92-07815-5
Sequence 5, Application PC/TUS9207815
GENERAL INFORMATION:
APPLICANT: DOWER, WILLIAM J.
APPLICANT: BARRETT, RONALD W.
APPLICANT: GALLOP, MARK A.
NEDELS, MICHAEL C
TITLE OF INVENTION: METHOD OF SYNTHESIZING DIVERSE
TITLE OF INVENTION: COLLECTIONS OF OLIGOMERS
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND
STREET: 1 MARKET PLAZA, STEWART TOWER, SUITE 2000
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/07815
FILING DATE: 19920916
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11509-36-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US92-07815-5

Query Match 100.0%; Score 11; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
|||||
DB 10 AGAGAGAGGAG 20

RESULT 8
US-09-147-915-6/c
Sequence 6, Application US/09147915A
Patent No. 6184034
GENERAL INFORMATION:
APPLICANT: Eastman, Alan
APPLICANT: Krieser, Ronald
TITLE OF INVENTION: Deoxyribonuclease II proteins and cDNAs
FILE REFERENCE: DC-0097
CURRENT APPLICATION NUMBER: US/09/147,915A
CURRENT FILING DATE: 1999-03-23
EARLIER APPLICATION NUMBER: PCT/US97/18262
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/028,539
EARLIER FILING DATE: 1996-10-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 30
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-147-915-6

Query Match 100.0%; Score 11; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
|||||
DB 29 AGAGAGAGGAG 19

RESULT 9
US-08-599-602-42/c
Sequence 42, Application US/08599602
Patent No. 5840839
GENERAL INFORMATION:
APPLICANT: WANG, R-F.; ROSENBERG, S.A.
TITLE OF INVENTION: ALTERNATIVE OPEN READING FRAME DNA OF A
TITLE OF INVENTION: NORMAL GENE AND A NOVEL HUMAN CANCER ANTIGEN ENCODED
TITLE OF INVENTION: THEREIN
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5-INCH, 1.44MB STORAGE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,602
FILING DATE: 09 FEB 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4209
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: cDNA
HYPOTHETICAL: YES
ANTI-SENSE: NO
US-08-599-602-42

Query Match 100.0%; Score 11; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
|||||
DB 29 AGAGAGAGGAG 19

RESULT 10
US-09-197-816-42/c
Sequence 42, Application US/09197816
Patent No. 6087110
GENERAL INFORMATION:
APPLICANT: WANG, R-F.; ROSENBERG, S.A.
TITLE OF INVENTION: ALTERNATIVE OPEN READING FRAME DNA OF A
TITLE OF INVENTION: NORMAL GENE AND A NOVEL HUMAN CANCER ANTIGEN ENCODED
TITLE OF INVENTION: THEREIN
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5-INCH, 1.44MB STORAGE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,816
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/599,602
FILING DATE: 09 FEB 1996
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4209
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
HYPOTHETICAL: YES
ANTI-SENSE: NO
US-09-197-816-42

Query Match 100.0%; Score 11; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||1111111111
DB 29 AGAGAGAGAG 19

RESULT 11
US-08-863-639A-26
Sequence 26, Application US/08863639A
Patent No. 5981185
GENERAL INFORMATION:
APPLICANT: Matson, Robert S.
APPLICANT: Coassin, Peter J.
APPLICANT: Rampal, Jang B.
APPLICANT: Caskey, C. T.
TITLE OF INVENTION: OLIGONUCLEOTIDE REPEAT ARRAYS
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheldon & Mak
STREET: 225 South Lake Avenue, 9th Floor
CITY: Pasadena
STATE: CA
COUNTRY: USA
ZIP: 91101
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Corel wordperfect 8 version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/863,639A
FILING DATE: May 28, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Joseph E. Muech
REGISTRATION NUMBER: 20,532
REFERENCE/DOCKET NUMBER: 11859-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (626) 796-4000

TELEFAX: (626) 795-6321
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid
US-08-863-639A-26

Query Match 100.0%; Score 11; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||1111111111
DB 18 AGAGAGAGAG 28

RESULT 12
US-08-560-313A-8
Sequence 8, Application US/08560313A
Patent No. 5763175
GENERAL INFORMATION:
APPLICANT: Sydney Brenner
TITLE OF INVENTION: Simultaneous Sequencing of Tagged Polynucleotides
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephen C. Macevicz, Lynx Therapeutics, Inc.
STREET: 3832 Bay Center Place
CITY: Hayward
STATE: California
COUNTRY: USA
ZIP: 94545
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: Power Macintosh
OPERATING SYSTEM: Macintosh OS ver. 7.5.2
SOFTWARE: Microsoft Word, vers. 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,313A
FILING DATE: 17-NOV-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Stephen C. Macevicz
REGISTRATION NUMBER: 30,285
REFERENCE/DOCKET NUMBER: sstllus
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 670-9365
TELEFAX: (510) 670-9302
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-560-313A-8

Query Match 100.0%; Score 11; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||1111111111
DB 13 AGAGAGAGAG 23

RESULT 13
US-08-611-155B-12
Sequence 12, Application US/08611155B

Patent No. 5780231
GENERAL INFORMATION:
APPLICANT: Sydney Brenner
TITLE OF INVENTION: DNA Extension and Analysis with Rolling Primers
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Stephen C. Macevicz, Lynx Therapeutics, Inc.
STREET: 3832 Bay Center Place
CITY: Hayward
STATE: California
COUNTRY: USA
ZIP: 94545
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows 3.1
SOFTWARE: Microsoft Word, vers. 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,155B
FILING DATE: 05-MAR-96
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/560,313
FILING DATE: 17-NOV-95
ATTORNEY/AGENT INFORMATION:
NAME: Stephen C. Macevicz
REGISTRATION NUMBER: 30,285
REFERENCE/DOCKET NUMBER: strplus
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 670-9365
TELEFAX: (510) 670-9302
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-611-155B-12
Query Match 100.0%; Score 11; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 1,1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11
DB 13 AGAGAGAGAG 23
RESULT 14
US-08-916-120A-14
Sequence 14, Application US/08916120A
Patent No. 5962228
GENERAL INFORMATION:
APPLICANT: Sydney Brenner
TITLE OF INVENTION: DNA Extension and Analysis with Rolling Primers
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Stephen C. Macevicz, Lynx Therapeutics, Inc.
STREET: 3832 Bay Center Place
CITY: Hayward
STATE: California
COUNTRY: USA
ZIP: 94545
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Microsoft Word, vers. 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,120A
FILING DATE: 22-AUG-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/611,155
FILING DATE: 05-MAR-96
ATTORNEY/AGENT INFORMATION:
NAME: Stephen C. Macevicz
REGISTRATION NUMBER: 30,285
REFERENCE/DOCKET NUMBER: 811-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 670-9365
TELEFAX: (510) 670-9302
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-916-120A-14
Query Match 100.0%; Score 11; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 1,1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11
DB 13 AGAGAGAGAG 23
RESULT 15
US-08-951-822-12
Sequence 12, Application US/08951822A
Patent No. 5989866
GENERAL INFORMATION:
APPLICANT: Deisher, Theresa A.
APPLICANT: Conklin, Darrell C.
APPLICANT: Raymond, Fenella
APPLICANT: Bukowski, Thomas R.
APPLICANT: Holderman, Susan D.
APPLICANT: Hansen, Birgit
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: NOVEL EGF HOMOLOGS
FILE REFERENCE: 96-20
CURRENT APPLICATION NUMBER: US/08/951,822A
CURRENT FILING DATE: 1997-10-16
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 63
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: oligonucleotide primer ZC15131
US-08-951-822-12
Query Match 100.0%; Score 11; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1,1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11
DB 17 AGAGAGAGAG 27
Search completed: July 28, 2003, 12:16:32
Job time: 1.6941 secs

Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 54.
 Location/Qualifiers
 source

1. 54
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U082M0282F14"
 /clone_1lb="Mouse 10kb plasmid U082M library"
 /sex="Female"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 BASE COUNT 1 a 21 c 5 g 27 t
 ORIGIN

Query Match 100.0%; Score 11; DB 17; Length 54;
 Best Local Similarity 100.0%; Pred. No. 5.4e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGAGAGAGAG 11
 |||||
 Db 15 AGAGAGAGAG 5

Search completed: July 28, 2003, 12:14:56
 Job time : 29.0947 secs

the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html

BASE COUNT 4 a 13 c 12 g 17 t

Query Match 100.0%; Score 11; DB 17; Length 46;
Best Local Similarity 100.0%; Pred. No. 5.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
DB 28 AGAGAGAGAG 18

RESULT 13
BF651086/c 52 bp mRNA linear EST 20-DEC-2000
LOCUS NF100H09EC1F1078 Elicited cell culture Medicago truncatula cDNA
DEFINITION clone NF100H09EC 5', mRNA sequence.

ACCESSION BF651086
VERSION BF651086
KEYWORDS EST.
SOURCE Barrel medic.
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 52)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research
JOURNAL Unpublished (2000)
COMMENT Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 52 Std Error: 0.00
Plate: 100 row: H column: 09
Seq primer: TCACACAGAGAAACAGCTATGAC.

FEATURES
SOURCE Location/Qualifiers

1. 52
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF100H09EC"
/clone_1lb="Elicited cell culture"
/tissue-type="Cell cultures derived from root tissues"
/dev-stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"
/note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

BASE COUNT 4 a 21 c 2 g 25 t

Query Match 100.0%; Score 11; DB 12; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.3e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
DB 43 AGAGAGAGAG 33

RESULT 14
BM395183/c 54 bp mRNA linear EST 17-JAN-2002
LOCUS BM395183

DEFINITION 50072-2-7-H10.f.2 Chilcoat/Turkewitz cDNA (large fraction)
Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION BM395183
VERSION BM395183.1 GI:18195236
KEYWORDS EST.

SOURCE Tetrahymena thermophila.
ORGANISM Tetrahymena thermophila
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Tetrahymenida; Tetrahymena.

REFERENCE 1 (bases 1 to 54)
AUTHORS Turkewitz, A.P., Karrer, K.M., Jahn, C., Orías, E., Kirk, K.E., Frankel, J., and Klobutcher, L.
TITLE EST from Tetrahymena thermophila, strain CU428.1, growing cells
JOURNAL Unpublished (2002)
COMMENT Contact: Turkewitz AP
Molecular Genetics and Cell Biology
University of Chicago
920 E. 58th Street, Chicago, IL 60637, USA
Tel: 773 702 4374
Fax: 773 702 3172
Email: apturkew@midway.uchicago.edu
Seq primer: T3.

FEATURES
SOURCE Location/Qualifiers

1. 54
/organism="Tetrahymena thermophila"
/strain="CU428.1"
/db_xref="taxon:5911"
/clone_1lb="Chilcoat/Turkewitz cDNA (large fraction)"
/note="Vector: Bluescript SK+; Details on library preparation can be found in Chilcoat and Turkewitz (2001) Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT 0 a 34 c 0 g 20 t

Query Match 100.0%; Score 11; DB 13; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.4e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
DB 44 AGAGAGAGAG 34

RESULT 15
A2996539/c 54 bp DNA linear GSS 27-APR-2001
LOCUS ZM0282F14R Mouse 10kb plasmid U06C2M library Mus musculus genomic
DEFINITION clone U06C2M0282F14 R, DNA sequence.

ACCESSION A2996539
VERSION A2996539.1 GI:13867766
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 54)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0282 row: F column: 14

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 40)
AUTHORS NCI/NIH-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: John Ensley, M.D.
 CDNA Library Preparation: Stratagene, Inc.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdpr/image/image.html

FEATURES
 source
 Trace considered overall poor quality
 Insert length: 1690 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 40
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="NCI-CGAP_HN3"
 /tissue_type="squamous cell carcinoma from base of tongue"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: tongue; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; cloned unidirectionally. Primer: 5' GAATTCGCGCAGAG 3' adaptor sequence: 5' (GA)TACTAGTCGAGTGTGTGTGTGTGTGTGT 3'")

BASE COUNT
 ORIGIN
 16 a 3 c 15 g 6 t

Query Match 100.0%; Score 11; DB 9; Length 40;
 Best Local Similarity 100.0%; Pred. No. 4.9e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
 |||||
 15 AGAGAGAGAG 25

Db 15 AGAGAGAGAG 25

RESULT 11
LOCUS B1252040 40 bp mRNA linear EST 17-JUN-2001
DEFINITION 602952695F1 NIH_MGC_100 Homo sapiens CDNA clone IMAGE:5087039 5', mRNA sequence.
ACCESSION B1252040
VERSION B1252040.1 GI:14802111
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 40)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>
 Plate: LLCM1843 row: m column: 24
 High quality sequence stop: 19.
 Location/Qualifiers
 1. 40
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="IMAGE:5087039"
 /tissue_type="NIH MGC 100"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT
 ORIGIN
 18 a 1 c 20 g 1 t

Query Match 100.0%; Score 11; DB 13; Length 40;
 Best Local Similarity 100.0%; Pred. No. 4.9e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
 |||||
 13 AGAGAGAGAG 23

Db 13 AGAGAGAGAG 23

RESULT 12
LOCUS BH812871/c 46 bp DNA linear GSS 02-MAY-2002
DEFINITION SALK_063356 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_063356, DNA sequence.
ACCESSION BH812871
VERSION BH812871.1 GI:20391326
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 46)
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J., and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
 Unpublished (2001)
 Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At5g40430.
 Class: TDNA tagged.
 Location/Qualifiers
 1. 46
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone_1lb="SALK_063356"
 /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at

ACCESSION	AA922878
VERSION	AA922878.1
	GI:3070187

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114)pb/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
7 a 1 c 17 g 4 t

Query Match 100.0%; Score 11; DB 17; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.4e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
Db 3 AGAGAGAGAG 13.

RESULT 6

AZ586653/c 33 bp DNA 1linear GSS 13-DEC-2000
LOCUS
DEFINITION
1M0392P11R Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCGIM0392P11 R, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS
AZ586653.1 GI:11708843

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 33)

REFERENCE
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

Unpublished (2000)

JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0392 row: P column: 11
Seq primer: CACACAGAAACAGCATATGACC
Class: plasmid ends
High quality sequence stop: 33.

FEATURES
source
1. 33
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCGIM0392P11"
/clone_1lb="Mouse 10kb plasmid UGCGIM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114)pb/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
2 a 19 c 0 g 12 t

Query Match 100.0%; Score 11; DB 17; Length 33;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
Db 23 AGAGAGAGAG 13

RESULT 7

AZ351541/c 36 bp DNA 1linear GSS 29-SEP-2000
LOCUS
DEFINITION
1M0089G11R Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCGIM0089G11 R, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS
AZ351541.1 GI:10430778

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 36)

REFERENCE
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

Unpublished (2000)

JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0089 row: G column: 11
Seq primer: CACACAGAAACAGCATATGACC
Class: plasmid ends
High quality sequence stop: 36.

FEATURES
source
1. 36
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCGIM0089G11"
/clone_1lb="Mouse 10kb plasmid UGCGIM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

```

TDNA.
Class: TDNA tagged.
location/Qualifiers
1. 24

```

```

BASE COUNT
ORIGIN
2 a 14 c 0 g 8 t
/cloae="SALK_019058.23.05.x"
/cloae_11b="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

```

Query Match	100.0%;	Score 11;	DB 17;	Length 24;
Best Local Similarity	100.0%;	Pred. No. 4.1e+04;		
Matches 11;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 AGAGAGAGGAG 11			
Db	14 AGAGAGAGGAG 4			

RESULT 4	AZ635993	25 bp	DNA	linear	GSS 13-DEC-2000
LOCUS	AZ635993				
DEFINITION	IM0493E20R Mouse 10kb plasmid U06C1M library Mus musculus genomic				
ACCESSION	AZ635993				
VERSION	AZ635993.1	GI:11758183			

SOURCE ORGANISM	house mouse. <i>Mus musculus</i>
REFERENCE 1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathii; Muridae; Murinae; Mus-1 (bases 1 to 25)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Petersen,T., Rellly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Nledernausern,A. and Wright,D.,Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL COMMENT	Unpublished (2000) Contact: Robert B. Weiss

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 1000 Std Error: 0.00
Plate: 0493 row: E column: 20
Seq primer: CACACAGAAACACACTATGACC
Class: plasmid ends
High quality sequence stop: 25.
Location/Qualifiers
1. .25
source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0493E20"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-
"
/note="Vector: pMD2.19; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
http://www.jax.org/resources/documents/dnares/". The DNA

```

Query Match	100.0%	Score 11:	DB 17:	Length 25:
Best Local Similarity	100.0%	Pred. No.	4.2e+04:	
Matches	11:	Conservative	0:	Mismatches 0:
				Indels 0:
				Gaps 0:
QY	1	AGAGAGAGAGAG	11	
Db	5	AGAGAGAGAGAG	15	

was hydrodynamically sheared by repeated passage through 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AR129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

RESULT	5
AZ992478	
LOCUS	AZ992478 29 bp DNA linear GSS 27-APR-2001
DEFINITION	ZM0277H07F Mouse 10kb plasmid U06C2M library Mus musculus genomic
ACCESSION	AZ992478 clone U06C2M0277H07 F, DNA sequence.
VERSION	AZ992478.1 GI:13863705

SOURCE	house mouse.
ORGANISM	Mus musculus.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus. 1 (bases 1 to 29)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Kelly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished (2000)
COMMENT	Contact: Robert B. Weiss

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel.: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0277 row: H column: 07
 Seq primer: CGTTGTAACGACGCCACAT
 Class: plasmid ends
 High quality sequence stop: 29.
 Location/Qualifiers
 1..29
 source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM2M027H07"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57Bl/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
http://www.jax.org/resources/documents/dnares/". The DNA

```

Plate: 0122 row: D column: 06
Seq primer: CGTTGTAATAACGACGCCACAGT
Class: plasmid ends
High quality sequence stop: 21.

FEATURES

Source

1. 21
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="tUGC1M0122D06"
/clone_lib="Mouse 10kb plasmid tUGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-gold, T1-resistant, F-"
/note="Vector: PWD42ny. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gII4732114[9b]Afl23072.1), a copy number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

1 a 7 c 0 g 13 f

Query Match	100.0%;	Score 11;	DB 17;	Length 21;
Best Local Similarity	100.0%;	Pred. No. 4e+04;		
Matches 11; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	AGAGAGAGGAG	1
Db	14	AGAGAGAGGAG	4

RESULT 2	
AZ861819	AZ861819
LOCUS	24 bp DNA
DEFINITION	2M0168L1R Mouse 10kb plasmid U06C1M library Mus musculus genomic clone U06C2M0168L2 R, DNA sequence.

REFERENCE
AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beecorn, T., Duval, B., Hamil, C.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Mus
1 (bases 1 to 24)

JOURNAL
COMMENT

University of Utah Genome Center
University of Utah
Rm. 300, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00

FEATURES

Source

1. 24
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2K018H12"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD44env: Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrolytically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD2 (gii4732114[gb]A12072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

9 a 0 c 15 g 0 t

Query Match	100.0%;	Score 11;	DB 17;	Length 24;
Best Local Similarity	100.0%;	Pred. No. 4.1e+04;		
Matches 11;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	AGAGAGAGGAG	11
Db	6	AGAGAGAGGAG	16

RESULT 3					
BH789331/c					
LOCUS					
DEFINITION					
	BH789331	24 bp	DNA	linear	GSS 02-APR-2001
	SALK_019058.23.05.x Arabidopsis thaliana cDNA insertion lines				
	Arabidopsis thaliana genomic clone SALK_019058.23.05.x, DNA sequence.				

REFERENCE
1 (bases 1 to 24)

JOURNA
COMMENT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 09:24:34 ; Search time 21.0947 Seconds
(without alignments)
8445.286 Million cell updates/sec

Title: US-09-902-772-1_COPY_650_660

Perfect score: 11

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapect 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_juv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	ID	Description
1	11	100.0	21	AZ371089	AZ371089 IM0122D06
2	11	100.0	24	AZ861819	AZ861819 2M0168L12
3	11	100.0	24	BH789331	BH789331 SALK_0190
4	11	100.0	25	AZ635993	AZ635993 IM0493E20
5	11	100.0	29	AZ992478	AZ992478 2M0277H07
6	11	100.0	33	AZ586653	AZ586653 IM0392P11

7	11	100.0	36	AZ351541	IM0089G11
8	11	100.0	37	BG744186	BG744186
9	11	100.0	38	BG684760	BG684760
10	11	100.0	40	AA922878	AA922878
11	11	100.0	40	BI252040	BI252040
12	11	100.0	46	BH812871	BH812871
13	11	100.0	52	BF651086	BF651086
14	11	100.0	54	BM395183	BM395183
15	11	100.0	54	AZ996539	AZ996539
16	11	100.0	56	AZ983943	AZ983943
17	11	100.0	57	AZ345621	AZ345621
18	11	100.0	59	AL767458	AL767458
19	11	100.0	60	BQ767486	BQ767486
20	11	100.0	61	AA503506	AA503506
21	11	100.0	63	BH857098	BH857098
22	11	100.0	68	BM521350	BM521350
23	11	100.0	68	AZ357658	AZ357658
24	11	100.0	69	CNS02YF4	CNS02YF4
25	11	100.0	75	N85423	N85423
26	11	100.0	75	AZ506280	AZ506280
27	11	100.0	76	BQ094017	BQ094017
28	11	100.0	76	AZ805134	AZ805134
29	11	100.0	79	AA668503	AA668503
30	11	100.0	79	AI082539	AI082539
31	11	100.0	79	AI085281	AI085281
32	11	100.0	79	AA142830	AA142830
33	11	100.0	79	AA620456	AA620456
34	11	100.0	79	AZ471783	AZ471783
35	11	100.0	79	AZ816527	AZ816527
36	11	100.0	79	L8315BX	L8315BX
37	11	100.0	81	BM301739	BM301739
38	11	100.0	81	AO073887	AO073887
39	11	100.0	84	BE324070	BE324070
40	11	100.0	85	AI152770	AI152770
41	11	100.0	85	AI438348	AI438348
42	11	100.0	85	BH225000	BH225000
43	11	100.0	85	CNS0255K	CNS0255K
44	11	100.0	87	BF718468	BF718468
45	11	100.0	88	AZ860469	AZ860469

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

21 bp DNA linear GSS 02-OCT-2000
IM0122D06F Mouse 10kb plasmid tucclm library Mus musculus genomic
clone tucclm0122D06 F, DNA sequence.
AZ371089
AZ371089.1 GI:10484789
GSS.
house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:09:50 ; Search time 3.03292 Seconds
(without alignments)
7482.227 Million cell updates/sec

Title: US-09-902-772-1_COPY_650_660

Perfect score: 11

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY_NUC

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCIT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	19	US-09-969-373-3925	Sequence 3925, Ap
2	11	100.0	20	US-10-215-112-9210	Sequence 9210, Ap
3	11	100.0	20	US-10-215-112-9325	Sequence 9325, Ap
4	11	100.0	20	US-10-215-112-9435	Sequence 9435, Ap
5	11	100.0	20	US-10-215-112-9437	Sequence 9437, Ap
6	11	100.0	20	US-10-215-112-9439	Sequence 9439, Ap
7	11	100.0	20	US-10-215-112-9715	Sequence 9715, Ap
8	11	100.0	20	US-10-215-112-9829	Sequence 9829, Ap
9	11	100.0	20	US-10-215-112-9938	Sequence 9938, Ap
10	11	100.0	20	US-10-215-112-9940	Sequence 9940, Ap
11	11	100.0	20	US-10-215-112-9942	Sequence 9942, Ap
12	11	100.0	24	US-09-852-980A-1	Sequence 1, Appl1
13	11	100.0	25	US-10-098-263B-10719	Sequence 10719, A
14	11	100.0	25	US-10-098-263B-10720	Sequence 10720, A
15	11	100.0	25	US-10-098-263B-22949	Sequence 22949, A
16	11	100.0	25	US-10-098-263B-63581	Sequence 63581, A

c	17	11	100.0	25	15	US-10-098-263B-80930	Sequence 80930, A
	18	11	100.0	25	15	US-10-098-263B-117135	Sequence 117135, A
	19	11	100.0	63	15	US-10-081-347-12	Sequence 12, Appl
	20	11	100.0	90	11	US-09-974-300-4057	Sequence 4057, Ap
	21	11	100.0	98	11	US-09-969-373-100	Sequence 100, App
	22	11	100.0	99	10	US-09-864-761-18548	Sequence 18548, A
	23	11	100.0	104	11	US-09-783-590-6469	Sequence 6469, A
	24	11	100.0	111	10	US-09-864-761-25653	Sequence 25653, A
	25	11	100.0	117	10	US-09-864-761-32279	Sequence 32279, A
c	26	11	100.0	120	11	US-09-969-373-1197	Sequence 1197, Ap
	27	11	100.0	124	10	US-09-864-761-17723	Sequence 17723, A
	28	11	100.0	129	11	US-09-728-444-210	Sequence 210, Appl
	29	11	100.0	141	15	US-10-081-347-19	Sequence 19, Appl
	30	11	100.0	146	11	US-09-728-444-211	Sequence 211, Appl
c	31	11	100.0	150	10	US-09-728-446-754	Sequence 754, Appl
	32	11	100.0	156	10	US-09-923-876-5331	Sequence 5331, Appl
c	33	11	100.0	162	11	US-09-878-574-11579	Sequence 11579, A
c	34	11	100.0	165	11	US-09-878-574-9778	Sequence 9778, Ap
	35	11	100.0	170	10	US-09-864-761-25721	Sequence 25721, A
	36	11	100.0	171	10	US-09-728-446-882	Sequence 882, Appl
c	37	11	100.0	174	11	US-09-764-847-1715	Sequence 1715, Appl
c	38	11	100.0	174	15	US-10-092-154-1715	Sequence 1715, Appl
c	39	11	100.0	176	11	US-09-764-847-1717	Sequence 1717, Ap
c	40	11	100.0	176	15	US-10-092-154-1717	Sequence 1717, Ap
c	41	11	100.0	177	10	US-09-770-696-299	Sequence 299, Appl
c	42	11	100.0	178	10	US-09-864-761-25409	Sequence 25409, A
c	43	11	100.0	180	10	US-09-294-093B-3061	Sequence 3061, Appl
c	44	11	100.0	182	15	US-10-085-906-159	Sequence 159, Appl
c	45	11	100.0	183	11	US-09-867-701-8586	Sequence 8586, Appl

ALIGNMENTS

RESULT 1
US-09-969-373-3925
Sequence 3925, Application US/09969373
Patent No. US20020133852A1
GENERAL INFORMATION:
APPLICANT: Efferitz, Roger J.
TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
FILE REFERENCE: 38-10(52679)A
CURRENT APPLICATION NUMBER: US/09/969,373
CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US 09/754,853
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: US 09/760,427
PRIOR FILING DATE: 2001-01-13
PRIOR APPLICATION NUMBER: US 09/855,768
PRIOR FILING DATE: 2001-05-15
NUMBER OF SEQ ID NOS: 4593
SEQ ID NO 3925
LENGTH: 19
TYPE: DNA
ORGANISM: Glycine max
US-09-969-373-3925
Query Match 100.0%; Score 11; DB 11; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Ox 1 AGAGAGAGGAG 11
Db 7 AGAGAGAGGAG 17
RESULT 2
US-10-215-112-9210
Sequence 9210, Application US/10215112
Publication No. US20030082596A1
GENERAL INFORMATION:
APPLICANT: Michael Miltmann

```

; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE OF INVENTION: Test3
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9210
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9210

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 3
US-10-215-112-9325
; Sequence 9325, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9325

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 4
US-10-215-112-9435
; Sequence 9435, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9435
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9435
```

```

US-10-215-112-9435

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 5
US-10-215-112-9437
; Sequence 9437, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9437
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9437

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 6
US-10-215-112-9439
; Sequence 9439, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9439
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9439

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 6 AGAGAGAGGAG 16

RESULT 7
US-10-215-112-9715/c
```

```
; Sequence 9715, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9715
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9715

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGAGAG 11
        |||
Db      15 AGAGAGAGAGAG 5

RESULT 8
US-10-215-112-9829/c
; Sequence 9829, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9829
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9829

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGAGAG 11
        |||
Db      15 AGAGAGAGAGAG 5

RESULT 9
US-10-215-112-9938/c
; Sequence 9938, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9938
; LENGTH: 20
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9938

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGAGAG 11
        |||
Db      15 AGAGAGAGAGAG 5

RESULT 10
US-10-215-112-9940/c
; Sequence 9940, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9940
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9940

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGAGAG 11
        |||
Db      15 AGAGAGAGAGAG 5

RESULT 11
US-10-215-112-9942/c
; Sequence 9942, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Miltmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9942
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-9942

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 20;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGAGAGAGAGAG 11
        |||
Db      15 AGAGAGAGAGAG 5
```

```
RESULT 12
US-09-852-980A-1
; Sequence 1, Application US/09852980A
; Patent No. US20020034752A1
; GENERAL INFORMATION:
; APPLICANT: ORDOVAS, JOSE M.
; APPLICANT: SCHAEFER, ERNST J.
; TITLE OF INVENTION: CPT TACIB POLYMORPHISM AS RISK FACTOR FOR DEVELOPMENT
; TITLE OF INVENTION: OF CORONARY HEART DISEASE
; FILE REFERENCE: TU-2004
; CURRENT APPLICATION NUMBER: US/09/852,980A
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,467
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-852-980A-1

Query Match
Best Local Similarity 100.0%; Score 11; DB 10; Length 24;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 10 AGAGAGAGGAG 20

RESULT 13
US-10-098-263B-10719
; Sequence 10719, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 10719
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-10719

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 1 AGAGAGAGGAG 11

RESULT 14
US-10-098-263B-10720
; Sequence 10720, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
```

```
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 10720
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-10720

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 1 AGAGAGAGGAG 11

RESULT 15
US-10-098-263B-22949
; Sequence 22949, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Miltman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 22949
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-22949

Query Match
Best Local Similarity 100.0%; Score 11; DB 15; Length 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
Db 8 AGAGAGAGGAG 18
```

Search completed: July 28, 2003, 12:23:26
Job time: 5.03292 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:39:19 ; Search time 4232.82 Seconds

(without alignments)
9948.861 Million cell updates/sec

Title: US-09-902-772-1

Sequence: 1 gaattccggaagcaataat.....gaaagaagcgcccaagaana 1447

Scoring table: .IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
GenBank:
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_dr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1447	100.0	1447	6 E31253	E31253 Protein hav
2	1409	97.4	1413	5 AY055661	AY055661 Gallus ga
3	1356	93.7	1528	6 E31254	E31254 Protein hav
4	1344	92.9	1516	5 GGERG	X77129 G.gallus ER
5	936	64.7	1808	10 AB031088	AB031088 Rattus no
6	921	63.6	3166	9 HUMERG2	M17254 Human erg2
7	889.6	61.5	2012	5 XLA74126	AJ224126 Xenopus l
8	881.6	60.9	2131	10 AB073078	AB073078 Mus muscu
9	867.2	59.9	2209	10 AB073080	AB073080 Mus muscu
10	828	57.2	2133	10 AB073079	AB073079 Mus muscu
11	801.8	55.4	1798	5 XLA74125	AJ224125 Xenopus l
12	760.2	52.5	3126	9 HUMERG11	M21535 Human erg p
13	571.2	39.5	1436	5 DRE249590	AJ249590 Danio rer
14	570.2	39.4	1729	10 MPEL11	X59421 Mouse Flt-1
15	563.8	39.0	2916	6 A36461	A36461 Sequence 2
16	563.8	39.0	2938	6 AR080101	AR080101 Sequence
17	563.8	39.0	2938	9 HSHMFL1	X67001 H.sapiens H
18	562.2	38.9	1359	9 AY029368	AY029368 Homo sapi
19	562.2	38.9	2394	9 BC010115	BC010115 Homo sapi
20	562.2	38.9	2957	9 HUMERGBFLI	M98833 Homo sapien
21	562.2	38.9	2959	9 BC001670	BC001670 Homo sapi
22	561.4	38.8	6849	12 AY029367	AY029367 Synthetic
23	560.6	38.7	1673	9 S45205	S45205 Flt-1-Frien
24	549.4	38.0	4403	5 XLF11G	X66979 X.laevis mR
25	518.6	35.8	1245	5 AF177538	AF177538 Danio rer
26	514	35.5	3490	5 CCNNAFL1	Y14773 Colurnix co
27	502.2	34.7	3545	5 CCEPLIONCO	Y14774 Colurnix co
28	493.2	34.1	762	9 S72621	S72621 EMS...erg l
29	478.2	33.0	1932	9 HUMFL1A	M93255 Human Flt-1
30	477	33.0	816	10 S66169	S66169 Erg-3-immun
31	387	26.7	14518	9 AP001423	AP001423 Homo sapi
32	387	26.7	100000	9 AP000021	AP000021 Homo sapi
33	387	26.7	100000	9 AP000163	AP000163 Homo sapi
34	387	26.7	340000	9 AP001731	AP001731 Homo sapi
35	373	25.8	588	9 S72865	S72865 EMS...EMS-e
36	358.6	24.8	1431	9 AF327066	AF327066 Homo sapi
37	320.4	22.1	66352	9 HSY17293	Y17293 Homo sapien
38	320.4	22.1	187131	2 AP001535	AP001535 Homo sapi
39	320.4	22.1	217382	9 AP001122	AP001122 Homo sapi
40	214.4	14.8	634	3 SULERG	M81067 Sea urchin
41	203	14.0	491	3 HD1311813	AJ311813 Hediste d
42	196.6	13.6	2180	3 AY060316	AY060316 Drosophila
43	180	12.4	1403	9 HSDNAFEV3	Y08979 H.sapiens F
44	180	12.4	1752	10 RN091679	Y091679 Rattus norv
45	180	12.4	169741	9 AC097468	AC097468 Homo sapi

ALIGNMENTS

```
RESULT 1
E31253
LOCUS      E31253
DEFINITION Protein having cell calcifying inhibitory activity and gene
ACCESSION E31253
VERSION   E31253.1 GI:13025685
KEYWORDS  JP 1999075871-A/1.
SOURCE    unidentifed.
ORGANISM  unidentifed.
REFERENCE 1 (bases 1 to 1447)
AUTHORS   Hiroyasu,I., Yoshinobu,H., Marjlio,P., Joel,R. and Helena,E.
TITLE      Protein having cell calcifying inhibitory activity and gene
encoding the same
```


Yeh, H., Abrams, W.R., Rosenbloom, J. and Pacifici, M.
Transcription factor ERG variants and functional diversification of
chondrocytes during limb long bone development
J. Cell Biol. 150 (1), 27-40 (2000)

JOURNAL MEDLINE
20351415
10893254
2 (bases 1 to 1413)
Iwamoto, M., Higuchi, Y., Enomoto-Iwamoto, M., Kurisu, K., Koyama, E.,
Yeh, H., Rosenbloom, J. and Pacifici, M.
The role of ERG (ets related gene) in cartilage development
Osteoarthritis. Cartil. 9 Suppl A, S41-S47 (2001)

JOURNAL MEDLINE
21535378
11680687
3 (bases 1 to 1413)
Iwamoto, M., Higuchi, Y., Yeh, H. and Pacifici, M.
Direct Submission
Submitted (05-DEC-2001) Oral Anatomy and Developmental Biology,
Osaka University Faculty of Dentistry, 1-8 Yamadaoka, Suita, Osaka
565-0871, Japan

FEATURES
source location/Qualifiers
1. 1413
/organism="Gallus gallus"
/db_xref="taxon:9031"
37. 1392
/note="alternatively spliced; transcription factor"
/codon_start=1
/product="erg isoform C-1-1"
/protein_id="FAL40889.1"
/db_xref="GI:17887441"
/translation="MASTIKREALSVSEDSLEFACVSPHLAKTEMTASSESSEYOT
SKMSPRVPOODMLSOPARTYIKMPCNPNVNGSRSPDCSVAKGKVVSSDNYGM
NYSYMEKHIIIPPNNMTNERRITVPADPLMTSDHYRQLEAVAYEGLPVDILF
QNDGKELKMTDQRLPSTNADILSHLYLKERGATLEFPMTSYPEVYOHIT
TRDPLVEDARSAWTSHPSTQSKATQSSSTVPKTEQDRLDYOILIGPSSILA
NPGSGOIOLOPFLIELSDSNGSCITWETNGEFTMDPDEVARWGEKRSKPNMY
DKLSRLRYYYDNIMTKVHGKRYAKFDHGLJAOLOPHPPESSMYKYPDLPYSS
YHHPQKMFVAPHPALPVTSSSFFAAPPYNSPTGCIYPRILPAHMPHSLCTY
Y"

BASE COUNT 424 a 370 c 311 g 308 t
ORIGIN

Query Match 97.4%; Score 1409; DB 5; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 TAGCAATTTAGCAATCAATTAATCTGTATCATTATGGCAAGCACTATTAGGAAGCA 86
DB 1 TAGCAATTTAGCAATCAATTAATCTGTATCATTATGGCAAGCACTATTAGGAAGCA 60

QY 87 TTATCAGTGTGAGTGAAGCAAGCAAGCTCTTGTGATGATGCTACGATGCCCCACCTT 146
DB 61 TTATCAGTGTGAGTGAAGCAAGCAAGCTCTTGTGATGATGCTACGATGCCCCACCTT 120

QY 147 GCAAAAGCAAAATGACAGCGCTCTCTCCAGTGAATATGGCAAAATCAAAAGATGAGC 206
DB 121 GCAAAAGCAAAATGACAGCGCTCTCTCTCCAGTGAATATGGCAAAATCAAAAGATGAGC 180

QY 207 CCGCGGCTCCCGCAGAGAGAGCTGTTATACACAGCCCCCGCCAGAGTTACATTAAAGATG 266
DB 181 CCGCGGCTCCCGCAGAGAGAGCTGTTATACACAGCCCCCGCCAGAGTTACATTAAAGATG 240

QY 267 GACTGTAAACCAACACAGCTTATGGGTCAAGAAATTCACCTATGACAGCGTGGCA 326
DB 241 GACTGTAAACCAACACAGCTTATGGGTCAAGAAATTCACCTATGACAGCGTGGCA 300

QY 327 AAAGAGAGAAATGTTAGCACTTACAGCAATGTTGGATGAAGTAACTATGGAAGCTACATG 386
DB 301 AAAGAGAGAAATGTTAGCACTTACAGCAATGTTGGATGAAGTAACTATGGAAGCTACATG 360

QY 387 GAAGAGAGCAATATTCGCGCTCCCAATATGACAAACAAATGAAAGAGTTATGTGCCA 446
DB 361 GAAGAGAGCAATATTCGCGCTCCCAATATGACAAACAAATGAAAGAGTTATGTGTGCCA 420

QY 447 GCAGATCTACGTATTAGAGACACAGCCATGTACGGCAGTGGCGTGGAGTGGCAGTGAAG 506
DB 421 GCAGATCTACGTATTAGAGACACAGCCATGTACGGCAGTGGCGTGGAGTGGCAGTGAAG 480

QY 507 GAGTATGTTCTCCAGACAGTGGACATCTTGTGTCAGAAATATGATGGAAAGAGTTG 566
DB 481 GAGTATGTTCTCCAGACAGTGGACATCTTGTGTCAGAAATATGATGGAAAGAGTTG 540

QY 567 TGTAAATGACCAAAAGATGACTTCACAGAGCTACAGCGGAGCTTAAGCAGATATCTTC 626
DB 541 TGTAAATGACCAAAAGATGACTTCACAGAGCTACAGCGGAGCTTAAGCAGATATCTTC 600

QY 627 CTGTACAGCTACAGTACCTCAGAGAGAGGAGCCACTTTATTTTCCAAATATCATCA 686
DB 601 CTGTACAGCTACAGTACCTCAGAGAGAGGAGCCACTTTATTTTCCAAATATCATCA 660

QY 687 GTTTACCAACAAACAGCAGCAAAATATACCAAGGCGAGATTTACCTTATGACAAGC 746
DB 661 GTTTACCAACAAACAGCAGCAAAATATACCAAGGCGAGATTTACCTTATGACAAGC 720

QY 747 AGGAGATCAGCGTGAAGAGTCAACAGCATCCACTCAGTCAAAAGCTACCAACCATCA 806
DB 721 AGGAGATCAGCGTGAAGAGTCAACAGCATCCACTCAGTCAAAAGCTACCAACCATCA 780

QY 807 TCTTCAAGATGCGCCAAACAGAAAGACACAGCTCCCTCAGTTAGATCTTATCAGATTCTT 866
DB 781 TCTTCAAGATGCGCCAAACAGAAAGACACAGCTCCCTCAGTTAGATCTTATCAGATTCTT 840

QY 867 GGACCAAGCAGCAGCGCTTGTGCAAAATCCAGGAGTGGGAGATATACAGTATGGCAGTTTC 926
DB 841 GGACCAAGCAGCAGCGCTTGTGCAAAATCCAGGAGTGGGAGATATACAGTATGGCAGTTTC 900

QY 927 CTACAGAGCTTGTGTGAGACAGCTCCCACTCACTGATCACTGCGAGGCGCACAAAT 986
DB 901 CTACAGAGCTTGTGTGAGACAGCTCCCACTCACTGATCACTGCGAGGCGCACAAAT 960

QY 987 GGGGAGTCAAGATGACAGACCGCTGATGAAGTGGCTGGGCTGGGAGAGAGAAAGC 1046
DB 961 GGGGAGTCAAGATGACAGACCGCTGATGAAGTGGCTGGGCTGGGAGAGAGAAAGC 1020

QY 1047 AAACCTAACATGAATATGACAAACTCAGCGGTGCACTTGTCTACTACTATGACAAAAAT 1106
DB 1021 AAACCTAACATGAATATGACAAACTCAGCGGTGCACTTGTCTACTACTATGACAAAAAT 1080

QY 1107 ATTATGACTTAAATGATGTTAAAGCTATGCTTACCAATTTGATTTCCAGGAAATGCTT 1166
DB 1081 ATTATGACTTAAATGATGTTAAAGCTATGCTTACCAATTTGATTTCCAGGAAATGCTT 1140

QY 1167 CAGGCGCTCAGGCTCACCCCTCAGAAATCATGATGATCAAAATACCATGACAGCTCCCG 1226
DB 1141 CAGGCGCTCAGGCTCACCCCTCAGAAATCATGATGATCAAAATACCATGACAGCTCCCG 1200

QY 1227 TACATGAGTTCCTACCATGACACACCCCGAGAAATGAACCTTTAGCTCCCATCCCTT 1286
DB 1201 TACATGAGTTCCTACCATGACACACCCCGAGAAATGAACCTTTAGCTCCCATCCCTT 1260

QY 1287 GCTTTCCCGGTAACTCATCCAGCTTTTGTGCTCCCTTAATCATCATGGAATTCACCA 1346
DB 1261 GCTTTCCCGGTAACTCATCCAGCTTTTGTGCTCCCTTAATCATCATGGAATTCACCA 1320

QY 1347 ACTGGAAGCATTCACCCCAATATACAGGCTGCGAGCTGCTCATATATGCTTCCCATTTGGC 1406
DB 1321 ACTGGAAGCATTCACCCCAATATACAGGCTGCGAGCTGCTCATATATGCTTCCCATTTGGC 1380

QY 1407 ACCTACTACTAAGTGGGAAAGAAAGAAA 1435
DB 1381 ACCTACTACTAAGTGGGAAAGAAAGAAA 1409

RESULT 3
E31254
LOCUS
DEFINITION Protein having cell calcifying inhibitory activity and gene

E31254 1528 bp DNA linear PAT 18-JUN-2001

RESULT 4
 GSERG LOCUS 1516 bp mRNA linear VRT 27-Apr-1995
 DEFINITION G.gallus ERG mRNA.
 X77159
 ACCESSION X77159.1 GI:790439
 VERSION ERG gene.
 KEYWORDS chicken.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 1516)
 Diderquin, P., Dewitte, F., Desbiens, X., Stehelin, D. and
 Duterque-Cogullaud, M.
 Mesodermal expression of the chicken erg gene associated with
 precartilaginous condensation and cartilage differentiation
 Mech. Dev. 50 (1), 17-28 (1995)
 JOURNAL 95329425
 MEDLINE 7605748
 PUBMED
 REFERENCE 2 (bases 1 to 1516)
 Duterque-Cogullaud, M.
 Direct Submission
 Submitted (17-JAN-1994) M. Duterque-Cogullaud, CNRS UA 1160,
 Oncologie Molculaire, Institut Pasteur, 1 rue Calmette, 59019
 Lille, FRANCE
 FEATURES
 source location/Qualifiers
 1. 1516
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /tissue_type="spleen"
 /clone_lib="lambda gt10"
 /dev_stage="adult"
 /gene="ERG"
 /gene="ERG"
 63. 1499
 /gene="ERG"
 63. 1499
 /codon_start=1
 /protein_id="CAA54404.1"
 /db_xref="GI:790440"
 /translation="MASTIKELSVSESDSEFCAYGSPHLAKTEMTASSSEYQOT
 SKMSPRPODMLSOPARYTIKMECPNPNVNSRSDSCVAKGKWSNDNGM
 NGSYMEKHIIPPMNTNTRVIVPADPTLMSTDHVROWLMAVVEYGLPDVILF
 QNIDKELCKMTKDDFORLPSYNADILSHLYLRTPLPHLTSDDVYKALONSRL
 MHRNKGATFIFPNTSVYPEARQRTTRDPLPEQARRAMTSHSHPTQSKATPDS
 STVKTEDDPQDLPQILGPTSSRLNPSGQIOLOMFLLELSDSSNRCITWGT
 NGEFKMTDPDEVARMRGERSKPMNMDKLSRLRYDDKINIMTKVHGKRYAKPFEH
 GIAOLOPHESSMYKPSDLPYMSYHNPQKMFVAPHPALPVTSSFFAALNP
 YMNSPGIYIPNTRIPAAHMPHLGTYT"
 BASE COUNT 458 a 392 c 327 g 339 t
 ORIGIN
 Query Match 92.9%; Score 1344; DB 5; Length 1516;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1433; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 241 CCCCGCCAGAGTTACCATTAAGATGAGTGTAAACCAACAGAGTTAATGGTCAAGA 300
 |||||
 Db 241 CCCCGCCAGAGTTACCATTAAGATGAGTGTAAACCAACAGAGTTAATGGTCAAGA 300
 QY 301 ATTACCTGATGACTGCGAGCGTGGCAAAAGAGGAAATGTTAGCATTGACACAATG 360
 |||||
 Db 301 ATTACCTGATGACTGCGAGCGTGGCAAAAGAGGAAATGTTAGCATTGACACAATG 360
 QY 361 TTGGGATGAACCTATGAGAGCTATGGAAGAGCAATATCCGCCCTCCAAATATGCAA 420
 |||||
 Db 361 TTGGGATGAACCTATGAGAGCTATGGAAGAGCAATATCCGCCCTCCAAATATGCAA 420
 QY 421 CCAATGAAGAGAGTATGTCGACAGATCCTACGTTATGAGACAGACATGTAC 480
 |||||
 Db 421 CCAATGAAGAGAGTATGTCGACAGATCCTACGTTATGAGACAGACATGTAC 480
 QY 481 GGCAGTGGCTGAGTGGGAGTGAAGAGATGATGTTCCAGACGTGACATGTTGT 540
 |||||
 Db 481 GGCAGTGGCTGAGTGGGAGTGAAGAGATGATGTTCCAGACGTGACATGTTGT 540
 QY 541 TCCAGAACATTTGATGGAAAGAGTTGTCTAAATGACCAAGATGACTTCCAGAGACTCA 600
 |||||
 Db 541 TCCAGAACATTTGATGGAAAGAGTTGTCTAAATGACCAAGATGACTTCCAGAGACTCA 600
 QY 601 CCGCGAGCTATAAGCGAGATATGCTCTGTCACACCTACACTACCTACAGAGAGA----- 654
 |||||
 Db 601 CCGCGAGCTATAAGCGAGATATGCTCTGTCACACCTACACTACCTACAGAGAGCTCTC 660
 QY 655 ----- 654
 |||||
 Db 655 ----- 654
 QY 661 TTCCACATTTGACTTCAGATGATGTTGATAAGGCTTACAAACCTCCACGGTTAATGC 720
 |||||
 Db 661 TTCCACATTTGACTTCAGATGATGTTGATAAGGCTTACAAACCTCCACGGTTAATGC 720
 QY 655 -----GAGGAGCCACTTTATTTTCCAAATATATCATGTTACCACAGAG 659
 |||||
 Db 655 -----GAGGAGCCACTTTATTTTCCAAATATATCATGTTACCACAGAG 659
 QY 721 ATGCTGAAACACAGAGGAGGAGCACTTTATTTTCCAAATATATCATGTTACCACAGAG 780
 |||||
 Db 721 ATGCTGAAACACAGAGGAGGAGCACTTTATTTTCCAAATATATCATGTTACCACAGAG 780
 QY 700 CAACGCAAGATATACCAACAAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 759
 |||||
 Db 700 CAACGCAAGATATACCAACAAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 759
 QY 781 CAACGCAAGATATACCAACAAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 840
 |||||
 Db 781 CAACGCAAGATATACCAACAAGGCGAGATTACCTTATGAGCAAGCAGAGATCAGCGT 840
 QY 760 GAGAGATCAGAGCATCCAGTCCAGTCAAGTCAAAAGCTTACCAACATCTTCAACAGTGC 819
 |||||
 Db 760 GAGAGATCAGAGCATCCAGTCCAGTCAAGTCAAAAGCTTACCAACATCTTCAACAGTGC 819
 QY 841 GAGAGATCAGAGCATCCAGTCCAGTCAAGTCAAAAGCTTACCAACATCTTCAACAGTGC 900
 |||||
 Db 841 GAGAGATCAGAGCATCCAGTCCAGTCAAGTCAAAAGCTTACCAACATCTTCAACAGTGC 900
 QY 820 CCAAAACAGAGACACAGCTCTCAGTTAGATCCTTATGAGATTTTGGACCAACAGCA 879
 |||||
 Db 820 CCAAAACAGAGACACAGCTCTCAGTTAGATCCTTATGAGATTTTGGACCAACAGCA 879
 QY 901 CCAAAACAGAGACACAGCTCTCAGTTAGATCCTTATGAGATTTTGGACCAACAGCA 960
 |||||
 Db 901 CCAAAACAGAGACACAGCTCTCAGTTAGATCCTTATGAGATTTTGGACCAACAGCA 960
 QY 880 GCGGCTTGGCAATCCAGGAGTGGGAGTACAGTATGGCAGTCTCTGAGAGCTTC 939
 |||||
 Db 880 GCGGCTTGGCAATCCAGGAGTGGGAGTACAGTATGGCAGTCTCTGAGAGCTTC 939
 QY 961 GCGGCTTGGCAATCCAGGAGTGGGAGTACAGTATGGCAGTCTCTGAGAGCTTC 1020
 |||||
 Db 961 GCGGCTTGGCAATCCAGGAGTGGGAGTACAGTATGGCAGTCTCTGAGAGCTTC 1020
 QY 940 TGTGCGACAGCTCCAACTCCAACTGCATCCAGTCCAGGAGGACAAATGGGAGTTCAAGA 999
 |||||
 Db 940 TGTGCGACAGCTCCAACTCCAACTGCATCCAGTCCAGGAGGACAAATGGGAGTTCAAGA 999
 QY 1021 TGTGCGACAGCTCCAACTCCAACTGCATCCAGTCCAGGAGGACAAATGGGAGTTCAAGA 1080
 |||||
 Db 1021 TGTGCGACAGCTCCAACTCCAACTGCATCCAGTCCAGGAGGACAAATGGGAGTTCAAGA 1080
 QY 1000 TGACAGACCTGATGAAAGTGGCTCGCGGTTGGGGAAGAGAGAAACCTTAACATGA 1059
 |||||
 Db 1000 TGACAGACCTGATGAAAGTGGCTCGCGGTTGGGGAAGAGAGAAACCTTAACATGA 1059
 QY 1081 TGACAGACCTGATGAAAGTGGCTCGCGGTTGGGGAAGAGAGAAACCTTAACATGA 1140
 |||||
 Db 1081 TGACAGACCTGATGAAAGTGGCTCGCGGTTGGGGAAGAGAGAAACCTTAACATGA 1140
 QY 1060 ACTATGACAACCTCAGCGCTGCACTTGCCTACTACTACTACTACTACTACTACTACT 1119
 |||||
 Db 1060 ACTATGACAACCTCAGCGCTGCACTTGCCTACTACTACTACTACTACTACTACTACT 1119
 QY 1141 ACTATGACAACCTCAGCGCTGCACTTGCCTACTACTACTACTACTACTACTACTACT 1200
 |||||
 Db 1141 ACTATGACAACCTCAGCGCTGCACTTGCCTACTACTACTACTACTACTACTACTACT 1200
 QY 1120 TTGATGTAAGAGTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCCCTCCAGC 1179
 |||||
 Db 1120 TTGATGTAAGAGTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCCCTCCAGC 1179
 QY 1201 TTGATGTAAGAGTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCCCTCCAGC 1260
 |||||
 Db 1201 TTGATGTAAGAGTATGCTTACAAATTTGATTTCCAGGAAATGCTCAGGCCCTCCAGC 1260
 QY 1180 CTCACCTCCAGATATCATCATATACCAATACCAATACCAATACCAATACCAATAC 1239
 |||||
 Db 1180 CTCACCTCCAGATATCATCATATACCAATACCAATACCAATACCAATACCAATAC 1239
 QY 1261 CTCACCTCCAGATATCATCATATACCAATACCAATACCAATACCAATACCAATAC 1320
 |||||
 Db 1261 CTCACCTCCAGATATCATCATATACCAATACCAATACCAATACCAATACCAATAC 1320
 QY 1240 ACCATGACACCCCGCAAGAGATGAGTGTATGCTCCCTGCTTCCCGGTAA 1299

Db	1321	ACCATGACACCCCCAGAGATGAAGTCTTGAGCTCCCATCCGCCCTGCTTGGCCGTAA	1380
QY	1300	CCTATATCAGCTTTTGTCTGCTCCCTTAATCATATCTGGAATTCACCAATGAGGACATCT	1359
Db	1381	CCTATATCAGCTTTTGTCTGCTCCCTTAATCATATCTGGAATTCACCAATGAGGACATCT	1440
QY	1360	ACCCCAATACAGAGCTGGCAGCTGCTCATATGCTTCCCATCTTGGGACCGACACATACAAAG	1419
Db	1441	ACCCCAATACAGAGCTGGCAGCTGCTCATATGCTTCCCATCTTGGGACCGACACATACAAAG	1500
QY	1420	TGGGGAAGAAGAAAGA 1435	
Db	1501	TGGGGAAGAAGAAAGA 1516	
RESULT 5			
AB031088			
LOCUS	AB031088	1808 bp	linear ROD 07-AUG-2001
DEFINITION	Rattus norvegicus VESPI4 mRNA for vascular endothelial cell specific protein 14, complete cds.		
ACCESSION	AB031088		
VERSION	AB031088.1	GI:15128488	
KEYWORDS	vascular endothelial cell specific protein 14.		
SOURCE	Rattus norvegicus liver cDNA to mRNA.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Rattus.		
REFERENCE	1 (bases 1 to 1808)		
AUTHORS	Aoki,T., Toyoda,H., Nishimoto,S., Tawara,T., Ukai,Y. and Komurasaki,T.		
TITLE	Identification of VESPI4, a vascular endothelial cell specific protein		
JOURNAL	Unpublished		
REFERENCE	Aoki,T., Toyoda,H., Nishimoto,S., Tawara,T., Ukai,Y. and Komurasaki,T.		
AUTHORS	Aoki,T., Toyoda,H., Nishimoto,S., Tawara,T., Ukai,Y. and Komurasaki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-AUG-1999) Tomomi Aoki, Taisho Pharmaceutical Co., Ltd., Medical Research Laboratories, Molecular Biology Lab.; 1-403, Yoshino-cho, Omiya, Saitama 3308530, Japan (E-mail:s17704@ccm.taisho.co.jp, Tel:048-663-1111, Fax:048-652-7254)		
FEATURES	Location/Qualifiers		
Source	1..1808		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	/tissue_type="liver"		
	1..1808		
	/gene="VESPI4"		
gene	63..1430		
	/gene="VESPI4"		
	/note="similar to human egr1"		
	/codon_start=1		
	/product="vascular endothelial cell specific protein 14"		
	/protein_id="BAB62744.1"		
	/dd_xref="GI:15128489"		
	/translation="MASTIKELALSVSKDLSFECAGYTPHLAKTEMTASSSSDYQQT SKSPRYPODDWL SOPPARVTIKMECPQSVNCSRNSEDSCYKGGKMGSPDTVM SYGSIYMEKHMPNMNTNRRVIVPADPLMSTDHVROWLENAVEYGLDVLVLE ONIDGKELCKMTKDDFORLPSPYNAIDILSLHYLRRTPLPHLITSDVDKALONSRL MHAANDLPEPRSRMTWGHSHPTPSKAAQSPSTVPTEDQRPIDPYQLGPTS SRLANPSGOIQLMQFLLELSDSSNSNCLTMEGTNEERKATPDDEVARWGRKSRP YMSYDTRALRYTDKNINMTKVHGKRYAKFDFHGLAQLQHPPESSLYKPSDLP NMNDYTHHPDKMNFVAPHPALPVTSSSFATNPENPWNSTPGIYPTNTRLPASHMPSH LGTYV"		
BASE COUNT	510 a	522 c	426 g 350 t
ORIGIN			
Query Match	64.7%	Score 936	DB 10; Length 1808;
Best local Similarity	80.2%	Pred. No. 5.8e-273;	
Matches 1140; Conservative	0;	Mismatches 270;	Indels 12; Gaps 3;

OY	8	GCAGACGAATATTTATTTATGACATTTATTTACGATTCATATATCTTGATACATATATGGC	67
Db	8	GAGAAAGCAATATTTATTTATTAACATTTTATTAACGATTCATATTAACCTTGATGCATATATGGC	67
OY	68	AAGCAGCTATTAAGAAGCATTATATCATGAGTGATGAAGACCAAGCTCTTTTGAAGTGTGC	127
Db	68	CAGCAGCTATTAAGAGGCGCTTGTCTGCTAGTTGTGTGACGAAGAACAGCTCATATTTAGAGTGTGC	127
OY	128	CTAGCGATGCGGCCCTTTGGCAAAACAGAAATGACAGCTCTCTCTTCAGAGAAATATGG	187
Db	128	CTTAGCGAAGCGCACACCTGGCTTAACAGAAATGACCGCATCTTTCCAGTGAATATGG	187
OY	188	GCAAAAGCTCAAAAGATGAGCCGCGCGCTTCCCGAGCAGACATGGTTATCACAGCCCCCGGC	247
Db	188	CCAGACATCCAAAGATGAGTCCCGAGGTCCACAAACAGAGACTGTGTCTGACACCCCGCAGC	247
OY	248	CAGAGTTACCATTTAAGATGAGGTGTAAACCCAAACAGGTAAATGGGCTCAAGGAATTCAC	307
Db	248	CAGGCTCACCTATCAAGATGAGATGGTGCACCTTATGCCAGGTGAACGGTCTCCAGGAATCCAC	307
OY	308	TGATGACATCGACGCGGCAAAAAGGAGGAAAAATGGTTATGACGTTACAGCAATATGGGAT	367
Db	308	TGAGCAATGCAAGTGTGACCAAAAGGGGAAAGTGTGGGACAGCCTGTGATCTGTGGGAT	367
OY	368	GAAGTATGGAAGCTACATGATGAAGAAAGCAATATTCGCCCTCCAAATATAGAACCAATATGA	427
Db	368	GAGCTAGCGGACGCTACATGAGAGAGAACACATGCCCGCCCAATATATGACACAGAAATGA	427
OY	428	ACGAAGATTTTGTGGCCAGCAGATCTTACGTTATGAGACACAGACCTATATACGGCAGTG	487
Db	428	ACGAGAGATGTATGCTCCCTGCAAGATCTTACATCTGTGGAGCACAGACCTATTCGGCAGTG	487
OY	488	GCTGAGAGTGGGACAGTGAAGAGATATGAGCTTCCAGACAGCTGGAACATCTGTTGTTCCAGAA	547
Db	488	GCTGAGTGGGACAGTGAAGAAATATACGGCTCTACAGCTGGACGCTTATATTTACAGAA	547
OY	548	CATTGATGGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCCAGAGACTTCACGCCGAG	607
Db	548	TATTGACGGGAAAGGAGTGTGTCAAGATACAAAAGATACCTTCCAGAGGCTCACCTCCGAG	607
OY	608	CTATTAACCAATATATCCCTGCTGACACTACATACACTCTCAGAGAGA-----GAGGAGCC	662
Db	608	CTACAATGCCACATCTCTCTCTCTCACTCTCAGTCTCAGAGAGACTCCCTCCCTTCACA	667
OY	663	ACTTTATTTTCCAAATATACATCAAGTTTACCCAGAA---GCAACGCAAGAAATTAACAAC	718
Db	668	TCTGACATCCGACGACGTTGATTAAGGCTTATCAAAAATCTCCACAGGTTAAATGACATGTAG	727
OY	719	AAGGCCAGATTTACTCTTATGAGCAAGCAGAGATCACCGTGTGAGATGCATACAGGCATCC	778
Db	728	AAACACAAATTTACTCTTATGAGCCCTCCAGAGATCAACCTGTGACCGGCGCACAGCAGCCACC	787
OY	779	CA---CTCAGTCAAAAAGCTACCAACACATCTTCAACAGATGCCCCAAACAGAAAGACCA	835
Db	788	CACCCCTCAGTCCAAAAGCTGCGCACAGCATCTCCCTCCACATGTGCCCAAAACCTGAAGACCA	847
OY	836	GCGTCTCAGTTAAGATCCCTTATCAAGATCTTGGACCGACGACGACGCGCTTTGCCAAATCC	895
Db	848	GCGTCTCAGTTAAGATCCCTTACAGATCTTGGAGCCACAGATGATGCCCTTTGGTAAATCC	907
OY	896	AGGAGATGGGAGATACAGCTATGTGGCAGTTCTCTCTGAGAGCTTGTGTGGGACAGCTCCAA	955
Db	908	AGGATAGTGGCAGATCCAGCTGTGTGGCAGTTCTCTGTAAACTCTCTGTGTGACAGCTCCAA	967
OY	956	CTCCAACTGCATCACTCTGGAGAGGCACAAATATGGGAGTTTCAAGATGACAGACCCCTGATGA	1015
Db	968	CTCCAACTGCATCACTCTGGAGAGGCACCAACAGGGGAGTTTCAAGATGACAGACCCGGATGA	1022
OY	1016	AGTGGCTCGCGCTTTGGGAGAGAGAAAAGCAAACTTAACATGAATATGACAAACTAG	1075
Db	1028	GGTGGCTCGCGCTTTGGGAGAGAGAAAAGCAAACTTAACATGAATATGACAAACTAG	1087

OY		1076	CCTGACACTGGTCACTAAGTCATAATGAACAATAATTATATGACTTAAGTCATGTGAACGCCTA	11353
OY		1088	CCGTGCGCCTCGCTACTACTACAGAAAAAAATCATATGACCAGAAGTGACGGGAAGCGCTA	11477
OY		1136	TGCTCACAAATTTGATTTGCCACGGAATGCTGCAGGCGCTCCAGCGCTCCAGCTCCAGAAATC	11959
Dd		1148	TGCTTACAGAATTTGACTTCCACGGGATTTGCCCAGGCGCTTGCAGAGCCCCATCCCCTCGAATC	12077
OY		1196	ATCCATGTACAAATATCCCATCAGACACTTCCCTTACATAGATGATTCACATGACACACGCCCA	12555
Dd		1208	GTCCTGTACAAAGTACACCCCTCCGACACTCCGTATCATGAGGCTTCATACACACCCACGCCCA	12677
OY		1256	GAAATGTAACATTGTAAGCTCCGCCCATCCCGCTGCTTTGGCCGCTAACCTATCCAGCTTTT	13151
Dd		1268	GAAATGTAACTTTTGtggctcccccacccctcgagccttcacactgtcatcttccaggtttgtt	13277
OY		1316	TGCTGCCCTATCATCTACTGGAATTGCACAACTGAGAGCATATACCCAATACCAAGCT	13757
Dd		1328	TGCTATCCCGCAACCATCTACTGGAATTTGCCGCAGCTGGGGGCTATCACCGGAACACTAGGCT	13877
OY		1376	GCCAGCTGCTCATATGCTTCCCATCTTTGGACACTACTACTA	1417
Dd		1388	CCGACCGCATATGCGCTCTCACACTGGGACCTACTACTA	1429
RESULT 6				
HOMERG2				
LOCUS	HUMERG2	3166 bp ss-RNA	linear	PRI 08-NOV-1994
DEFINITION	Human erg2 gene encoding erg2 protein, complete cds.			
ACCESSION	M17254			
VERSION	M17254.1 GI:182186			
KEYWORDS	erg 2 protein.			
SOURCE	Human cell line COLO 320, cDNA to mRNA, clone lambda 12.			
ORGANISM	Homo sapiens.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 3166)			
AUTHORS	Rao,V.N., Papas,T.S. and Reddy,E.S.			
TITLE	ery, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation			
JOURNAL	Science 237 (4815), 635-639 (1987)			
MEDLINE	87263429			
PUBMED	3299708			
FEATURES				
source	Location/Qualifiers			
	1..3166			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="21q22.3"			
gene	257..1645			
	/gene="ERG"			
CDS	257..1645			
	/gene="ERG"			
	/note="erg 2 protein"			
	/codon_start=1			
	/protein_id="AAA52398.1"			
	/db_xref="GI:182187"			
	/db_xref="GeneDB:G00-119-884"			
	/translation="MIOVPDPAHIKEALSVSEDOISLEPCAYGTPHLAKTEMTASS			
	SDSDGQTSKMSRPYPQQDMLSQPAPRTIKECNPISOYNGSRNSPDECSVAKSGKMVG			
	SPDTVSNTGSYEEMKHPPPNMTNTERVIVPADPTLMSTDHYRQMELMAVKRGYLGL			
	DNVLILFQNIDGKLCKTKDDFORLIPSYNADLLSHLAYLETPLPHLDVIDVDVA			
	LONSPRLMANANDLPYEPERRSAMTHGHETPOSKAAPSPSTVPEKDORPOLDEPY			
	KIOLPPTSRLANPGSOILOQLLELIDSNSGSIWEGTNGEFNEDDEVARARM			
	GERSKRNPMNDXKSRLRALRYDYDKIMTYKGRAVAKFDPHGIAOLAOPHPRESSIX			
	KYPSDLPIYMGSVIHAHPQKMNFVAHPALPVLTSSFTFAARPYNPSPTGCIYPTRLI			
	TSNHPSHLGITY"			
	291..298			
	/note="8-dp repeat sequence"			
	509..516			
	/note="8-bp repeat sequence"			
	3143..3149			
	/note="putative"			
BASE COUNT	915 a 770 c 768 g 713 t			

ORIGIN	227 bp upstream of PvuII site; chromosome 21.
Query Match	63.6%; Score 921; DB 9; Length 3166;
Best Local Similarity	81.1%; Pred. No. 2,2e-268;
Matches 1100:	Conservative 0; Mismatches 245; Indels 12; Gaps 2;
OY	74 TATTAGACACATTTATCAGTGTGAGTGGAAGACCAGTCCTTGTGGTGTGAGTGTGCCACGG 133
Dd	289 TATCAAGAACCCCTTATCATGTTGTGAGTAGAGAGACCGACGTCGTTGTTGATGTGTGCCTACGG 348
OY	134 ATCGCCCCACCTTGCACAAGACAGAATAATCACAGCCCTCCCTTCACAGTAAATTTGGGCAAC 193
Dd	349 AACCGCACACCTGTGCTTAAGACAGAGATBACCGGCTCTCTCTCCACAGCATTTGACAGAC 408
OY	194 ATCAAAGATGAGCCCGCGCTTCCCACAGACAGCTGTATTACAGACCCCGCGCAGAGT 253
Dd	409 TTCCAAGATGAGCCACCGCTCTCCACAGAGATTTGGCTGCTCAACCCCGACGAGGT 468
OY	254 TACCATTTAAGTGTGAGTGTAAACCCACAGGTTAATGGGTGCMAGGAATTCACCTGATGA 313
Dd	469 CACCATCAAAATGGAAATGTAACCTTAGCAGAGTGAATGGCTCAAGGAATCTCCCTGATGA 528
OY	314 CTGCAGCGCTGGCAAAAAGAGGAAAAATGGTTAGCGTTCAACACAAATGTTGGGATTAACCTA 373
Dd	529 ATGCAGTGTGCCCAAGAGCGGGAAAGTGTGTGGCGACGCCACAGACCGTTGGGAGATGAAC 588
OY	374 TGGAAGCTTACATGGAAGAGAAAGCATATTTCCGCTCTCAATATATGACAAACCAATGAAAG 433
Dd	589 CGGCAAGCTACATGGAGAGGAAGACATGCCACCCCAAAACATGACACAGAACGAGCGAG 648
OY	434 AGTATTTTGTCCAGCAGATCTCTACGTTATYGAGCAGACAGACCATGTACGGCAGTGGCTGGA 493
Dd	649 AGTTATCTGTCGAGAGATCTCTAGCGCTATGAGATPACAGACCATGTGCGCGAGTGGCTGGA 708
OY	494 GTGGGCGATGGAAGAGTATGCTCTTCCAGAGSTGGACACTCTGTTGTTCCAGAACATTTGA 553
Dd	709 GTGGGCGGTGAAGAGATATGGCGCTTCCAGAGCTCAACATCTTGTATTCCAGAACATCTGA 768
OY	554 TGGGAAGAAGTGTGTATAAATGACCAAAAGATGACTTCCAGAGACTCACGCCAGCTATAA 613
Dd	769 TGGGAAGSAACCTGTGTCAAGATGACCAAGAGAGACTTCCAGAGGCTCACGCCAGCTACAA 828
OY	614 CGCAGATATCTCTCTGTACACACTTACCTACCTAGAGAGA-----GAGGAGCCAG 664
Dd	829 CGCGGACATCTCTCTCTCACATCTCCACTACCTCAGAGAGACTCTCTTCCACATTTTGC 888
OY	665 TTTTATTTTCCAAATNATCAGTATTNCCGAGAAGCAACGCAAGAAATAAACAAGGCC 724
Dd	889 TTTCAGATGATATTGTATTAAGCTTTACAAACATCTTCCACAGGTTAATGATGTATGAAACAC 948
OY	725 AGATTTCACCTATAGCAGCAAGCGAGAGATCAGCTGTGACGAGTACACAGCCATATCCA---C 781
Dd	949 AGATTTCACATATGAGCCCCCAGAGATACAGCTGTGACCGGTACAGGCCACCCACGCC 1008
OY	782 TCAGTCAAAAAGCTACCCCAACCATCATCTTCAAGAGTGGCCCAAAACAGAAAGCACAGCGTCC 841
Dd	1009 CCAGTCGAAAGCTGTCTCAACCATCTCTTCCACAGTGGCCAAAACGTGAAGACACAGCGTCC 1068
OY	842 TCAGTTAGATCTTATCAGATTTCTTGGACCGACACAGCAGCCGCTGTGCAATTCAGAGGAG 901
Dd	1069 TCAGTTAGATCTTATCAGATTTCTTGGACCAACAGTAAGCCGCTGTGCAATTCAGAGGAG 1128
OY	902 TGGGCAAGTACAGTATGAGGAGTTCTATCTGAGAGCTTGTGCGACAGCTCCAACTCCAA 961
Dd	1129 TGGCCAGATCCAGCTTTTGGCAGTTCTCTGTGAGGCTCTGTGTGGGACAGCTCCAACTCCAG 1188
OY	962 CTCGATCACTCTGGGAGGCGCACAAATTGGGGAATTTCAAGATGACAGACCTGTATGAAGTGGC 1021
Dd	1189 CTGATATACCTCTGGGAAGGCACCAACGGGGAATTTCAAGATGACAGGATCCCGACGAGTGGC 1248
OY	1022 TCGCGGTTTGGGAGAGGAAAAAGCAAACTTACATGACTATATACAAACTCAGCGCTGC 1081
Dd	1249 CCGCGCGCTGGGAGAGGCGAAAGACCAAACTTCAATGACTATGACTATGACTACGCGCGC 1308

QY	1082	ACTTCGCTACGACATGACCAAAATTTATGACATTAAGTTGATGTAAAGCCTATGCCTA	1141
Db	1309	CTCCTCGTTACTGATGATGACAAGAATCATATGACCAAGGTCATGAGGAGCGCTACGCTA	1368
QY	1142	CAAAATTTGATTTCCACGGGAATGCGTCAGAGCCCTTCAGGCTCAGGCTCAGATCATCCAT	1201
Db	1369	CAAGTTGACGCTTCACAGGGATGGCGCAGAGCCCTTCAGAGCCACCCCGGAGTGATCTCT	1428
QY	1202	GTCAAAATACCCATGACAGCTCCCTTCATGATGATTTCTTACCATGACACACCCCAAGAGAT	1261
Db	1429	GTACAAGATACCCCTCAGACCTCCGCTACATGGCTCCGATATCAGCCACACCAAGAGAT	1488
QY	1262	GAAGTTTGATCTCCCATCCCGCTGGTTGGCCGCTTAACCTCATCCAGGCTTTTGGCTGC	1321
Db	1489	GAACTTTGTGGCGCCCAACCTCCAGGCTCCCGGTGACATCTTCCAGTTTGTGGCTGC	1548
QY	1322	CCCTAATTCATCTGTAATTCACCACTGAGAGCATGACCCCAATACAGAGCTGCAGC	1381
Db	1549	CCCAAGCCCATCTGGAATTCACCACTGAGGCGGTATATACCCCAACTAGGCTCCCGAC	1608
QY	1382	TGCTCATATGCTTCCCATCTTGGCAGCTACTACTAA	1418
Db	1609	CAGCCATATGCTTCTCATCTGGGCGACTACTACTAA	1645
RESULT 7			
LOCUS	XLAJ4126	2012 bp	mrna linear VRT 21-JAN-2000
DEFINITION	Xenopus laevis erg gene (erg-E).		
ACCESSION	AJ224126		
VERSION	AJ224126.1 GI:5420047		
KEYWORDS	ERG gene; transcription factor.		
SOURCE	Xenopus laevis.		
ORGANISM	Xenopus laevis.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.		
AUTHORS	1 (bases 1 to 2012)		
TITLE	Baltzinger M., Mager-Heckel, A.M. and Remy, P.		
JOURNAL	XI erg: expression pattern and overexpression during development		
MEDLINE	plead. for a role in endothelial cell differentiation		
PUBMED	Dev. Dyn. 216 (4-5), 420-433 (1999)		
REFERENCE	10633861		
AUTHORS	2 (bases 1 to 2012)		
TITLE	Baltzinger M.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (24-FEB-1998) Baltzinger M., UPR9005, MMDCD, Centre		
REFERENCE	National de la Recherche scientifique, 15, RUE Rene Descartes,		
TITLE	67084, FRANCE		
JOURNAL	Location/Qualifiers		
FEATURES	1..2012		
SOURCE	/organism="Xenopus laevis"		
	/db_xref="taxon:8335"		
	/clone="lambda ZAPII-erg E"		
	/clone_11d="lambda ZAPII"		
	/dev_stage="stage 30 embryo"		
	225..1595		
gene	/gene="erg"		
	225..1595		
CDS	/gene="erg"		
	/function="Transcription factor"		
	/codon_start=1		
	/protein_id="CA64567.1"		
	/db_xref="GI:5420048"		
	/db_xref="SPTREMBL:Q9W629"		
	/translation="MIDTPDPSAHIKELAVSSEDSLFECTYGTPIHLTKTEMTASS		
	SSYDGCKSKSPRVPOODMLQSPSPRYTIKMCNPSQVNGSRSPDSCSIGKSKMG		
	GSDVGVNGVSYIEEKHISPPNMTNRRTVADPTIMSTDHVRQMLFWAIEYGLP		
	DVYVLLNQNDKRELCKMTKEDRHRLPTSPNADILSLHTLYLBERGASLFFPSSVYQ		
	DANORIPSRDULSTEPSRSAMWNHAPAPSKASQPTTVKTEDDPRQADPYILGAT		
	SSRIANGPSQIQIOWFLLEILSDSSNSNCTIWEGTNGEPMKMDPDEVARWGERSK		

BASE COUNT	634 a	479 c	395 g	504 t	
ORIGIN	<p> PNMNDRKLSRLRYDYDINIMTYVHGKRYAKRPFDPHGLAOLQHPPESTPYKXPSLE PYMSYXHAFHPOKMNFVAHPHPALREVTVSSFFPAANAATVWNSPTGSIYNTRLRPAHSMSS HLGYTY" </p>				
Query Match	61.5%: Score 889.6; DB 5; Length 2012;				
Best Local Similarity	79.7%: Pred. No. 7,6e-259;				
Matches 1077;	Conservative 0; Mismatches 269; Indels 6; Gaps 2;				
QY	74	TATTTAAGGAAGCATTATCATGTGGTGAAGTGAAGACCAAGCTCTTGTGGTGGTGCCTACGG	133		
DB	257	TATCAAAAGAGCTCTTTCAGTGGTGAAGTGAAGATCAGCTCTTATTCGAGTGCACCTACGG	316		
QY	134	ATGCGCCACCTTGGAAAGACAGAAATACAGCCCTCCCTCCAGTGAATTTGGGCAAC	193		
DB	317	AAGCGACATCTTACTTAACGGAGATACCGCATCTTCTCCAGTGAAGTGAAGGCAAC	376		
QY	194	ATCAAAATGAGCCCGCGCTTCCCGACAGAGCTGTTATCACAGCCCCGGCGAGAGT	253		
DB	377	CTCGAAATGATGATCCACGTGTGCTTCAACAAGATGGCTCTCACAGCTTCATCAGAGGT	436		
QY	254	TACCATTTAAGTGGGTATACCAACACAGTTAATGGGCAAGGAAATTCACCTGATGA	313		
DB	437	GACCATCAAGATGATCAACCTAGTCAAGTCAATGATCAAGAGACTCCCGGATGA	496		
QY	314	CTGCACCTGGCAAAAGAGGAGAAATGGTTAGCAGTTCCAGACATGTTGGGATGACTA	373		
DB	497	CTGCAGTATAGGAAGAAAGAGACAAATGGGTGAGGGTGGGATATGTTGGAAATGAATTA	556		
QY	374	TGGAAGCTACATGGAGAAGACATATTCGCGCTCCAAATATGACAACCAATGACAGAG	433		
DB	557	TGGCAGCTACATAGAAAGAAACACATTTCTCCCAACAACTGACCAATGAGAGAG	616		
QY	434	AGTTATTTGGCAGAGATCTTACGTTATGGAGACAGACCATATGACGGCATGGCTGGA	493		
DB	617	AGTATTTTACTCTGTACCTTACTTTTGGAGACAGATATATGTACAGACATGGCTAGA	676		
QY	494	GTGGGACGTGAAGAGATATGGTCTTCCAGACGTGGACATCTTGTGTTCCAGAACTTGA	553		
DB	677	GTGGGCATTAAGAAATATACGGCTTCCAGATGTGATGTGCTCTGTTTCCAAACATGCA	736		
QY	554	TGGGAABAGTGTGGTAAATGACCAAAAGATGATCTCCAGATCTCAAGCCGCGACTTAA	613		
DB	737	CGGCAAGGAATTTATGCAGATGACCAAGAAAGAAATTTCCATGACTCAAGCCCAACTCAAA	796		
QY	614	CGCAGATATCTCTGTCTCACACCTACACTTACTCTAGAGAGAGAGAGCCACTTTATTTT	673		
DB	797	TGCTGATATCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	856		
QY	674	TCCAAATTCATCAGTTTACCCAGAGACAGCAAAAGATTAACCAAGGCCAGATTAC	733		
DB	857	CCCAAACTCATCTCGTTTATCAAGACGCAAAAGAAATTCCTCCAGCGCAAGGCAAGTTTATC	916		
QY	734	TTATGAGCAACGAGAGATCAGGTGGAGTGCAGTACAGCCATCCACTAGTCAAAAGC	793		
DB	917	ATATGAGGCCCTCAAGAGATCTGGTTGACAAATCA---TCTCTACACACATCAAAAGC	973		
QY	794	TACCAACCATCATCTTCAACAGTGGCCAAAACGAAAGACCAGCTCTCTAGTTAGATCC	853		
DB	974	ATCTCAAC---CATCTACACAGTTTCCCAAAACGAAAGACCCACAGGCCACAGTATGATCC	1030		
QY	854	TTATTCAGATCTTGGAGCGACACAGCGCTTGGCAAAATCCAGAGGATGGGCGAGATACA	913		
DB	1031	TTATTCAGATCTTGGGCGCAACACAGCGCTTGGCAAAATCCAGAGGATGGGCGAGATACA	1090		
QY	914	GCTATGGAGTCTTCTAGTGAAGCTTCTGTGGAGAGCTCCAACTCCAACTGATCACTCTG	973		
DB	1091	GCTCTGGGAATTTTACTGGAAATCGCTTTGGGATGTTTCCAACTCCAACTGATCACTCTG	1150		
QY	974	GGAGGGCAAAATGGGAGTTCAAGATGACAGACCTGATGAAGTGGCTGGCGCTTGGGG	1033		
DB	1151	GGAGGAAACCAATGAGAAATTAAGATGACCGATCTCGATGAAGTGGCAAGACTTGGGG	1210		

QY	1034	AGAGAGAAAAAGCAAACTTAACATGACATATGACAAACTAGACCTGTACACTGGCTACTA	1093
Db	1211	GGAGAGAAAAAGCAAACTTAACATGACATATGACAAACTAGACCTGTACACTGGCTACTA	1270
QY	1094	CTATGACAAAAATATATATGACTTAAAGTTTCATGATGTAAGCGTATGCCATCAAAATTTGATT	1153
Db	1271	CTACGATAAAAATATATATGACTTAAAGTTTCATGATGTAAGCGGCTATGCCATCAAAATTTGATT	1330
QY	1154	CCAGGAAATGGCTTCAGGCCCTCCAGCCTTCAGCCTTCAGAAATCATATCAATATACCC	1213
Db	1331	TCATGGGATGGCTTCAGGCCCTTCAGCCTTCAGCCTTCAGAAATCAATCATATGATCAAGTATCC	1390
QY	1214	ATCAGACCTCCCTCCATGAGTGTCTTCACATGATGATCCACACCCCGAGAAATGAACTTTGATGC	1273
Db	1391	CTCAGAACCTTCATATATATGAGCTATATACATGACATCCACAGAAATGAACTTTGATGC	1450
QY	1274	TCCCATTCGCCCTGCTTTGGCCCGTAATCCATATCCAGCTTTTGTGCTCCCTTAATCCATA	1333
Db	1451	CCCCCATTCGCCGACCTTTACCTGTGACATCGTCAAGTTCTTTGACAGCCCTTAATGCATA	1510
QY	1334	CTGGAATTCACCAACTGGAGGATCTACCCCAATATCCAGGCTGGCAGCTGCTATATGCC	1393
Db	1511	CTGGAATTCACCAACTGGAGATATTTATCCAAATATCTGGCTGGCAGCTGACCATATATGTC	1570
QY	1394	TTCCCATCTTGGGACCTACTACTATGAGGGGA	1425
Db	1571	TTCTCATCTTGGAGACCTACTATTTAAATCAAGA	1602

	RESULT 8
LOCUS	AB073078
DEFINITION	AB073078 Mus musculus Erg mRNA, mouse homolog of Human ets-related gene Erg,
ACCESSION	AB073078
VERSION	AB073078.1 GI:16191716
KEYWORDS	Mus musculus
SOURCE	Mus musculus cDNA to mRNA, clone:Erg-1.
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Oawara,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and Sakaki,Y. Mus musculus Erg mRNA Published in Database (2001) 2 (bases 1 to 2131) Oawara,R., Noguchi,H., Taylor,T.D., Takeda,T., Hattori,M. and Sakaki,Y. Direct Submission Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-Chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattorigsc.riken.go.jp, URL:http://bgp.gsc.riken.go.jp/, Tel:-81-45-503-9111, Fax:-81-45-503-9170)
TITLE	JOURNAL
FEATURES	source
1..2131	/organism="Mus musculus"
/db_xref="taxon:10090"	/chromosome="16"
/clone="Erg-1"	1..2131
/gene="Erg"	148..1536
/note="homolog of human ets-related gene (ERG), transcript variant 1"	/codon_start=-1
	/protein_id="BAB69948.1"
	/db_xref="GI:16191717"
	/translation="MTQIVDPAAHKEALSVSEDSLEFCAYGPPLAKTEMTASS SSDYGGTSKSPRPQDDWLSQAPARVTKMECPNSOVNCSRMSPECSYNGCKRWVG

Query Match	Best Local Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps	2;
60.9%;	79.3%;	881.6;	10;	2131;						
Query Similarity	Pred. No.	2,1e-256;								
Matches 1075;	Conservative	0;	Mismatches	269;	Indels	12;	Gaps	2;		
BASE COUNT	582 a	590 c	530 g	429 t						
ORIGIN										
Query	74	TATTAAAGAACATTATCACTGAGTGTGAAGACCAAGCTCTTTGTTAGTGTGCTTACGG	133							
Db	180	TATTAAAGAGAGCCTTTCAGTGTGTGAGACGAGAGCACTATTTATGATGTGCTTACGG	239							
Qy	134	ATGCCCCCACTTGTGCAAAAGACAGAAATGACAGCCTCTCTTCAGTGAATATGGGCAAC	193							
Db	240	AACGCCACACTGGCTTAAGACAGAGATGACCGCATCTTCCAGTACATATGGCCAGAC	299							
Qy	194	ATCAAAAGATGAGCCCCGGCTTCCCGACAGAGATGTGTTATCACAGCCCCGGCCAGAGT	253							
Db	300	ATCCAAAGTATGCTCCAGAGATGCTCTCAGCAGAGACTGGCTGTCTTAAGCCCCAGCCAGGT	359							
Qy	254	TACCATTAAGATGAGTGTAAACCCAAACAGAGTTAATGGGTCAAGGAATTCACCTGATGA	313							
Db	360	CACCATCAAGATGAGTGTAAACCCATGACAGGTGAATGGTTCCAGGAATCACCCTGATGA	419							
Qy	314	CTGCAGCGTGGCAAAAGAGGAGAAATGGTTAGCAGATTACAGACAATGTTGGGATGAACCTA	373							
Db	420	GTCGAGTGTGAACAAAGGTGGGAGATGTGGGAGCCCGGATCTGTGGGGAATGACCTA	479							
Qy	374	TGGAAGCTACATGGAAGAGAAAGATATTCGCCCTCCCAATATGACACCAATGAAGCAAG	433							
Db	480	CGGCAGCTACATGGAAGAGAAAGATGTGGCGCTCCCAATATGACACCAATGAAGCGCAG	539							
Qy	434	AGTATTGTGCGACAGATCCTACGTTATGAGACACAGACCATTAGCGAGTGGCTGGA	493							
Db	540	AGTATGTGCTCTGACAGATCCTACTGTGTGAGACACAGACCATTGCTCGACAGTGGCTGA	599							
Qy	494	GTCGGCAGTGAAGAGATATGCTCTTCCACAGCGTGGACATCTTGTGTTCCAGAACTTGA	553							
Db	600	GTCGGCGGTGAAGAAATATGGCTCTTCGATGTGAGAGTGTCTTACTATTATGAAATATGCA	659							
Qy	554	TGGGAAGAGTGTGTAAATATGACCAAAAGATGATCTTCCAGAGACTCAGCGGCACTTAA	613							
Db	660	TGGGAAGAGTGTGTGCAAGATGACAAAGAGATGATCTTCCAGAGGCTCAGCGGCACTTAA	719							
Qy	614	CGCAGATATTCCTCTGTCACACCTACACTACCTCAGAGAGA-----GAGGAGCGAC	664							
Db	720	TGCGGACATCTTCTTCCACATCTCCACTACCTTAGAGAGAGCTCCCTTCCACATCTGAC	779							
Qy	665	TTTTTATTTTCCAAATACATCAGTTTACCCAGAAAGATGATCTTCCAGAGGCT	724							
Db	780	TTTCGAGATGAGTGTATTAAGGCTTTACAAAAATCTCCACGGTTATATGATGACGAAACAC	839							
Qy	725	AGATTATACCTTATGAGACAGAGAGAGATACGCTGGAACGATCACAGCCATC---CCAC	781							
Db	840	AGATTATACCTTATGAGCTCTCCAGAGATACAGCTGTGACCGGCGCACAGCCACTCACCCC	889							
Qy	782	TCAGTCAAAAGCTTACCCCAACATCATCTTCCACAGTGTCCCAAAACAGAACCAAGCGTCC	841							
Db	900	TCAGTCCAAAGCTGTCTAGCCATCTTCCCTTGTGACGTGCCCAAAACTGTAAACCAAGCGTCC	959							
Qy	842	TCAGTTAGATCTTATCAGATTTCTTGGACCGACAGCAGCGCTTGTGCAATATCCAGGAG	901							
Db	960	TCAGTTAGATCTTATCAGATTCGTTGGACCGACAGCAGCGCTTGTCTAATCCAGGTAG	1019							
Qy	902	TGGGCAAGATACAGCTATGCGAGTTCTTACTGTGAGCTTCTGTGGACAGCTCCAACTGCA	961							
Db	1020	TGGGCAAGATACAGCTGTGGCAGTTCCTGCTGGAATCTCTTCCAGACAGCTCCAACTTCCA	1079							

QY 962 CTCATCACCCTGGAGGACCAAAATGGGAGATTCAAGATGACAGACCCTGATGAATGGC 1021
 |||||
 Db 1080 CTCGATACCTGGGAAGGACCAACGSGGAGTTCAATATACAGACCCGGACAGGTGGC 1139
 QY 1022 TCGGCGTGGGAGAGAGAAACCAAACTTAACATGACTATGACAACTCAGCGTGC 1081
 |||||
 Db 1140 TCGGCGTGGGAGAGAGAAACCAAACTTAACATGACTATGACAACTCAGCGCGC 1199
 QY 1082 ACTTCGCTACTATGACAAATAATATGACTAAAGTCAAGTGAAGCGTATGGCTTA 1141
 |||||
 Db 1200 CTCGCGTACTATGACAAATAATATGACTAAAGTCAAGTGAAGCGTATGGCTTA 1259
 QY 1142 CAAATTTGATTTCCAGGAAATCCCTCAGGCCCTCCAGCCCTCAGCCCTCAGAAATCATTCAT 1201
 |||||
 Db 1260 CAAATTTGATTTCCAGGAAATCCCTCAGGCCCTCCAGCCCTCAGCCCTCAGAAATCATTCAT 1319
 QY 1202 GTACAAATACCATCAGACCTCCCTTACATGATGCTTCTACATGACACCCCGAGAAAT 1261
 |||||
 Db 1320 GTACAAATACCATCAGACCTCCCTTACATGATGCTTCTACATGACACCCCGAGAAAT 1379
 QY 1262 GAATTTGTAGTCCCATCCCTCCCTGCTTCCCGTACCTCAATCCAGCTTTTGTGTC 1321
 |||||
 Db 1380 GAATTTGTAGTCCCATCCCTCCCTGCTTCCCGTACCTCAATCCAGCTTTTGTGTC 1439
 QY 1322 CCCTAATCCTACTGGAATTCACCAACTGAGGATCTACCCCAATACAGGCTGCGCAGC 1381
 |||||
 Db 1440 CCCTAATCCTACTGGAATTCACCAACTGAGGATCTACCCCAATACAGGCTGCGCAGC 1499
 QY 1382 TGCTATATGCTTCCATCTTTGGCAGCTTACTACTA 1417
 |||||
 Db 1500 CAGCATATGCTTCCATCTTTGGCAGCTTACTACTA 1535

RESULT 9
 AB073080 2209 bp mRNA linear ROD 17-OCT-2001
 LOCUS
 DEFINITION Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,
 transcript variant 3, complete cds.
 ACCESSION AB073080
 VERSION AB073080.1 GI:16197544
 KEYWORDS
 SOURCE Mus musculus cDNA to mRNA, clone:Erg-3.
 ORGANISM Mus musculus

REFERENCE
 1. Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.
 Mus musculus Erg mRNA
 TITLE Published Only in Database (2001)
 JOURNAL 2 (bases 1 to 2209)
 AUTHORS Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.

TITLE Direct Submission
 JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical
 REFERENCE and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 AUTHORS 1-7-22 Suenho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gs.c.riken.go.jp, URL:http://nrg.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)

FEATURES
 Location/Qualifiers

1..2209
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"
 /clone="Erg-3"
 1..2209
 /gene="Erg"
 148..1608
 /gene="Erg"
 /note="homolog of human ets-related gene (ERG), transcript
 variant 3"
 /codon_start=1

/protein_id="Bab69950.1"
 /db_xref="GI:16197545"
 /translation="MIOTVDPAAHIKREALSVSEDOLEFECAYGPHLAKTEMTASS
 SDPYGKSWSPRPVPOODLSLOPARTIMEQNPSONVNSRSPESCVNKGKWKV
 SPTPVGKSYKMEKHVPPNNATNERRIVYADPNLTDBYRQMLBNAVYEGIL
 DVDVLEFONIDGELCKMTKDQRLRSLNADILSHLHLETPHILTSDDVDA
 LONSPLMRHNRNTGGAFFEPNTSVYPEAQRITTRDLDYEPPRSAMTGHSLTPQ
 SKAOPSAPVPRKEDQRPQDLYQILGYSRLANPDSGQIOLMOFLLELSDSSNS
 NCTWEGTNGEERKSPKPNMYDKLSALRYTDKNTMTYVHGR
 YAKPDPHGIAOLAOPHPRESSLYKYPQSDLPYAGSVYHARHKNFVSPHPALPVTS
 SFPSAPRPVNSPTGTYPNTRIPASIMPHSLGTY"
 BASE COUNT 606 a 608 c 545 g 450 t
 ORIGIN
 Query Match 59.9%; Score 867.2; DB 10; Length 2209;
 Best Local Similarity 77.8%; Pred. No. 5e-252;
 Matches 1111; Conservative 0; Mismatches 233; Indels 84; Gaps 2;
 QY 74 TATTAAAGAAAGCATTTATCACTGAGTGAAGAACCACTCTTGTGAGTGTGCTTACGG 133
 |||||
 Db 180 TATTAAAGAGAGCTTGTCACTGTGAGCGAGAGCAAGTACTATTTGAGTGTGCTTACGG 239
 QY 134 ATGGCCCACTTGGCAAAAGACAGAAATGACAGCTCTCTTCCAGTGAATATGCGCAAC 193
 |||||
 Db 240 AACGCCACCTGGCTAAGACAGAGATGACCGCATCTCTTCCAGTGAATATGCGCAAC 299
 QY 194 ATCAAGATGAGCCGCGGCTTCCCGCAGAGACTGGTTATCAAGAGCCCGGCGCAGAGT 253
 |||||
 Db 300 ATCCAAATAGTGTCCAGAGATCCCTCAGCAGAGATGGCTGTCTCAAGCCCGCAGAGT 359
 QY 254 TACCATTAAAGATGAGTGAATACCCAAACAGGTTAATGGTCAAGGAATTCACCTGATGA 313
 |||||
 Db 360 CACCATCAATGATGAGAGTCAACCTTACAGTGAATGTTCCAGGAATCACCTGATGA 419
 QY 314 CTCGAGCGTGGCAAAAGGAGGAAATGTTTACAGTTCAGACAAATGTTGGATGAACAT 373
 |||||
 Db 420 GTGAGTGTGAACAAAGGTGGAGATGTGGGACCGGAGTATGTGGGATGAGCTA 479
 QY 374 TGGAAAGCTACATGGAAGAGAGATATTCGCGCTCCCAATATGACAAACCAATGAACGAG 433
 |||||
 Db 480 CGGAGCTACATGGAAGAGAGATATTCGCGCTCCCAATATGACAAACCAATGAACGAG 539
 QY 434 AGTTATTTGCCACAGACATCTTACGTTATGAGACACAGACCATGATGAGGCTGGA 493
 |||||
 Db 540 AGTATGATGTCCTCGACAGATCTACTGTGTGAGACACAGACCATGATGAGGCTGGA 599
 QY 494 GTGGGACGGAAGAGATGAGTCTTCCAGAGTGGAGATGTTGTTGTCAGAAACATGA 553
 |||||
 Db 600 GTGGGCGGTGAAGAAATATGCGCTCTCGATGTGAGAGTCTTACTATTTAGAAATATCGA 659
 QY 554 TGGGAAAGAGTGTGTTAAATGACCAAGATGACTTCCAGAGACTCAGCGGAGCTATAA 613
 |||||
 Db 660 TGGGAAAGAGCTGTGCAAGATGACAAAGATGACTTCCAGCGGCTACCGCGAGCTACAA 719
 QY 614 CGCAGATATCTCTGTACACCTTACACTACCTCAGAGAGA----- 654
 |||||
 Db 720 TGGCAGCATTTCTCTCACATCTCCACTACCTCAGAGAGACTCCCTTCCACATCTGAC 779
 QY 655 ----- 654
 Db 780 TTCCGATGACGTTGATTAAGGCTTTACAAAACCTCCACGAGTTAATGATGCCAANAAC 839
 QY 655 --GAGAGCCACTTTATTTTTCAAAATACATGATGTTTACCGAGAACGCAAGAAAT 712
 |||||
 Db 840 AGGGGGTGACAGCTTTATTTTCCAAATATCTTACATATCCGGAAGCTACGCAAGAAAT 899
 QY 713 AACCAAGGCGAGATTTTACCTTATGAGACAGGAGAGATCAGCGTGAAGTACAG 772
 |||||
 Db 900 TACAACATGAGCGAGATTTTACCTTATGAGCTCCAGAGATGACCTCGAGCCGCGACAG 959
 QY 773 CCATC---CCACGAGTCAAAAGCTTCCACAGCATCATCTTCAAAAGTGGCCCAAAACAG 829
 |||||
 Db 960 CCACCTCACCCCTGAGTCAAAAGCTGCTGACATCTCTCTGACAGTGCACAAACTGA 1019

QY 830 AGACAGCGTCCCTAGTATGCTTATGAGATTCCTTGAGCCGACGACGCGCTTGC 889
 DB 1020 AGACAGCGTCCCTAGTATGCTTATGAGATTCCTTGAGCCGACGACGCGCTTGC 1079
 QY 890 AATCCAGGAGTGGGCGAGATACAGCTATGAGCTTCTAGCTGAGCTTCTGTCGAGAC 949
 DB 1080 TAATCCAGGATGAGCGAGATCCAGCTGTGGCACTTCTGCTGCACTCTCTGTACAGAC 1139
 QY 950 CTCCAACTCCAACTGATCATCTGAGGAGGCGCAATATGGAGTTCAAGTACAGACCC 1009
 DB 1140 CTCCAACTCCAACTGATCATCTGAGGAGGCGCAATATGGAGTTCAAGTACAGACCC 1199
 QY 1010 TGATGAGTGGCTGGGGGTTGGGAGAGAGAGAAAGCAACCTTACATGAACTGTGACAA 1069
 DB 1200 GGAGAGGTGGCTGGGGGTTGGGAGAGAGAGAAAGCAACCTTACATGAACTGTGACAA 1259
 QY 1070 ACTGAGCGGTGACTGCTGCTACTGATGACAAATAATATTATGACTAAAGTCTAGTAA 1129
 DB 1260 GCTGAGCGGGCGCTCGCTACTGATGACAAATAATATTATGACTAAAGTCTAGTAA 1319
 QY 1130 ACCGTATGCTCAAAATTTGATTTCCAGGAAATGCTGAGCGCTCCAGCCTCACCTCC 1189
 DB 1320 GCGCTAGCCTCAAAATTTGATTTCCAGGAAATGCTGAGCGCTCCAGCCTCACCTCC 1379
 QY 1190 AGATCATGCTGATGACAAATATCCATGACAGCTCCCTGATGATGATGATGATGACAA 1249
 DB 1380 TGAATCCTGCTGATGACAAATATCCATGACAGCTCCCTGATGATGATGATGATGACAA 1439
 QY 1250 CCCCAGAGATGAAATTTGATGCTCCCAATCCCTGCTGTTGGCCCTTAACCTTACCTCAG 1309
 DB 1440 CCCCAGAGATGAAATTTGATGCTCCCAATCCCTGCTGTTGGCCCTTAACCTTACCTCAG 1499
 QY 1310 CTTTTCCTGCTGCTGCTTATGCTGATGAAATTTACCACTGAGAGGCTTACCCCAATAC 1369
 DB 1500 TTTTCCTGCTGCTGCTTATGCTGATGAAATTTACCACTGAGAGGCTTACCCCAATAC 1559
 QY 1370 CAGGCTGCAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417
 DB 1560 TAGGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1607

RESULT 10
 AB073079 2133 bp mRNA linear ROD 17-OCT-2001
 LOCUS
 DEFINITION Mus musculus Ery mRNA, mouse homolog of Human ets-related gene ERG,
 transcript variant 2, complete cds.
 ACCESSION AB073079.1 GI:16197542
 VERSION
 KEYWORDS
 SOURCE Mus musculus cdna to mRNA, clone: Ery-2.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Oawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.
 TITLE Mus musculus Ery mRNA
 JOURNAL Published only in Database (2001)
 REFERENCE 2 (bases 1 to 2133)
 AUTHORS Oawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Sphiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gs.c.riken.go.jp, URL: http://ngp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 FEATURES
 SOURCE 1..2133
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"

/clone="Erg-2"
 1..2133
 /gene="Erg"
 149..1540
 CDS
 /gene="Erg"
 /note="homolog of human ets-related gene (ERG), transcript variant 2"
 /codon_start=1
 /protein_id="BAB6949.1"
 /translation="MIQTPDPAHIEALSVSEDSQLEFECAGVPHLAKTENTASS
 SSDYGTSSKSPRYPOODMLSOAPARTIMECNBSQVNSRNSPDECSYNGKGMV
 SPDTGMSVGSYMEKEHPNMTTNRRTVADPTLMSDHRWLFEAVYEGLL
 DVLVLEFQNDKELEKMTDDQRLPTSTNADILSHLYLRETPPLPHLTSDDVDA
 LQNSPRLMHRNTGGAAVIFPNTSVYPEATQRTTTPAQPSPSAVKTEDQRLD
 YQLGPTSSRLANPDSQIOLMOFLLELSDSSNSCIMEGNGEKMDPDVAR
 WGEBSKSPNNMYDKLSRALRYDYDKNMTVEHGRVRYKDFDGLAOLAPHPRESSL
 KYKPSDLPYMSYHAHPQKNMFVSPHPALPVYSSPFAFPNPNYMTSGTIVPNTRL
 PASHMPSHLCTTY"
 BASE COUNT 587 a 579 c 531 g 436 t
 ORIGIN
 Query Match 57.2%; Score 828; DB 10; Length 2133;
 Best Local Similarity 76.8%; Pred. No. 4.2e-240;
 Matches 1046; Conservative 0; Mismatches 295; Indels 21; Gaps 2;
 QY 74 TATTAGGAACATTATCAGTGTGAGTGAAGACCACTCCTGTTGATGTCCTACGG 133
 DB 181 TATTAGGAGGCTTGTGACGTTGTGAGCGAGGACCACTCATATTGATGTCCTACGG 240
 QY 134 ATGCCCCACCTTGCAAGAGAGAGAAATGACAGCCTCCTTCCAGTAAATGAGCAAC 193
 DB 241 AAGGCCACACTGGCTAAGACAGAGATGACCGCATCTCTTCCAGTACTATGCGCAGAC 300
 QY 194 ATCAAGATGAGCCGCGCTTCCAGCAGAGACTGGTTATACAGGCCCGCGCAGAGT 253
 DB 301 ATCCAAATGATGATCCAGAGTCCCTCAGCAGAGACTGGCTGTCTCAAGCCCGCAGAGG 360
 QY 254 TACCATTAAATGAGTGTAAACCAACAGGTTAATGGGTCAAGAGAAATCAGCTGATGA 313
 DB 361 CACCATTAAATGAGTGTAAACCAACAGGTTAATGGGTCAAGAGAAATCAGCTGATGA 420
 QY 314 CTGAGCGTGGCAAAAGGAGAAATGCTTAGAGTTGACAGACATTTGGATGAGACTA 373
 DB 421 GTGCACTGTGAACAAAGGTGGAGATGTGGGCGACCCGGATRACGTGGGATGACTA 480
 QY 374 TGGAACTACATGGAAGAGAGATATTCGCTCCAAATATGACCAACATGAAAGCAG 433
 DB 481 CGGAGCTACATGGAAGAGAGATATTCGCTCCCAATATGACCAACATGAAAGCAG 540
 QY 434 AGTTATTTGCCAGCAGATCTTACGTTATGAGAGCAGACCATGTACGGCAGTGGTGA 493
 DB 541 AGTATGCTCCGCAATATCTTACGTTATGAGAGCAGACCATGTACGGCAGTGGTGA 600
 QY 494 GTGGGCGAGTGAAGAGTATGCTTCCAGAGTGGAGATCTTGTGTTCCAGAACTTGA 553
 DB 601 GTGGGCGGTAAGAAATATGAGCTCTCCGAGTGGAGTGTCTTATTTAGATATACGA 660
 QY 554 TGGGAAAGAGTGTGTAATATGACCAAGATGACTTCCAGAGCTCAGCGGAGCTATAA 613
 DB 661 TGGGAAAGAGCTGTCAAGATGACAAAGATGACTTCCAGCGGCTCAGCGGAGCTCAA 720
 QY 614 CGCAGATATCTCTGTCACACCTACACTACCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 673
 DB 721 TGGCGAGATTTCTCTGCAATCTCCACATCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 674 TCCAAATATCATGATTATCCAGAGCA-----CGCAAGAAATAC 715
 DB 781 TTCCGATGAGCTGTATGAGCTTTACAAACTCTCCAGGTTATGATGATCCAGAAACAC 840
 QY 716 AACAGGCCAGATTTATCATGAGCAAGCGAGAGATCAGCGTGAAGAGTACAGACCA 775
 DB 841 AGGGGGGCGAGCTTTATTTTCCAAATATCTTACATGATATTCGGAAGCTACGCAAGAAAT 900

QY	655	-GAGGACCACTTTTATTTTCCAAATATACATACAGTTTACCCAGACGACGCAAGAT	712
Db	1019	AGGAGTCCCTCTTTTATTTTCCAACTCTCCGTTTATCAAGACGCAACCAAGAT	1078
QY	713	AACAACAAGGCGAGATTTTACCTTATAGACAAGCGAGATCAGCGTAGAGTACAG	772
Db	1079	CCCCGACGAGCAAGATTATATCATATGAGCCCTCAAGAGATCTGTTGGACAATCAT	-
QY	773	CCATGCCACTAGTCAAAAGTACCCAACTATCTTCAACAGTGGCCCAAAACAGANA	832
Db	1137	-CTGGACACACCATCAAAAGCATCTCAAC---CATCTCCACAGTGTCCCAACAGAA	1192
QY	833	CCAGCGTCCCTAGTTAGTTCCTTTACATTTCTTGGACCGACGACGACCGCTTGC	892
Db	1193	CCCCAGGCGCAGACTAGTATGATCTTATCAGATTTCTTGGCCCAACAGACGCACTT	1255
QY	893	TCCAGGAGTGGGAGATACAGTATGAGCACTTCTACTGAGGCTTCTGTGGACAGCT	952
Db	1253	TCCAGGAGTGGGAGATTCAGCTCTGTGGCAATTCCTTACTGGAAGCTTGTGGAT	1312
QY	953	CAACTCCAACTGCAATCCTGGGAGGCGACAATTTGGGAGTTCAAGATGACAGACCT	1012
Db	1313	CAACTCCAACTGCAATTCCTTGGGAGGCAACCAATGGAATTTAAGATGACGATCC	1372
QY	1013	TGAAGTGGCTGGGTTGGGAGAGAGAAAGCAAACTTAACATGAACTATGACAACT	1072
Db	1373	TGAAGTGGCAAGAGCTTGGGGGAGAGAAAGCAACCACTGAACATATGACAACT	1432
QY	1073	CAGCCGTGCACTTGCTACTATATGACAAATATTTATGACTAAAGTTCATGGTAA	1132
Db	1433	CAGCCGTGCACTTGCTACTACTAGATAAATATTTATGACTAAAGTTCATGGCA	1492
QY	1133	CTATGCGCTACAAATTTGATTTCCAGGAAATGGCTCAGGCCCTCCAGCTCACC	1192
Db	1493	CTATGCTTACAAATTTGATTTTCATGGATTTGCTCAGGCTCTTCAACTCATCT	1552
QY	1193	ATCATTCATGTACAAATACCCTATAGACCTCCCTACATGATGCTCTACATGCA	1252
Db	1553	ATCATTCATGTACAAATATCCCTCAGAACTTCATATATGAGTCATATCATGCA	1612
QY	1253	CCACAAGATGAACTTTGATAGTCCCATCCCTCGCTTGGCCGTAACCTCATCG	1312
Db	1613	ACAAGAAGTGAATTTGTTGGCCCCCATCCCGCACTTTACTCTGATCATGTCA	1672
QY	1313	TTTGTGCGCCCTTATCATCTGTGAATTCACAACTGAGGCAATACCCCATTA	1372
Db	1673	CTTTGCGACCCCTTATGATCATCTGGAATTCACCAACTGGAAGTATTTATCCA	1732
QY	1373	GCTGCGACGCTGCTATATGCTTCCCATCTTTGGCACTTACTACTAAGTGGG	1425
Db	1733	GCTGCGACGCTGCAATATGATGTCTTCTCATCTTGGAACTACTATTAATCA	1785
RESULT 12			
HUMERGL1			
LOCUS	HUMERGL1	3126 bp	mRNA
DEFINITION	Human erg protein (ets-related gene)	mRNA, complete cds.	
ACCESSION	M21535	M17390	
VERSION	M21535.1	GI:182182	
KEYWORDS	erg protein.		
SEGMENT	1 of 2		
SOURCE	Human, cell line COLO 320, cDNA to mRNA, lambda-7.		
ORGANISM	Homo sapiens		
REFERENCE	Emukoyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Reddy,E.S., Rao,V.N. and Papas,T.S.		
TITLE	The erg gene: a human gene related to the ets oncogene		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	84 (17), 6131-6135	(1987)
MEDLINE	87317608		
PUBMED	3476934		
FEATURES	Location/Qualifiers		

	source	1..3126	/organism="Homo sapiens"
		/db_xref="taxon:9606"	
CDS		195..1286	
		/note="erg1 protein"	
		/codon_start=1	
		/protein_id="AA35811.1"	
		/db_xref="GI:182185"	
		/translation="TWGSDPDTVGNMNYSSYMEKHPPEPMNMTINRRVLPADPLTMSI DHVKOMLEMAVKEYGLDPVNILLFNONIDKRELCKTKRDPRLRPSTYAADLISLIHLH LREIPLEPHLISDDVDQALONSPRLMHANRDLTPLEPPRSKMTHGHFTPOSKAOP PSYPTPEDQRPDLPYOILGPSTSRLANPSSGOIQMWLFLELSDSNSSCTITWE TINGEGRKDPDEVARMRWGERKSKEPNMNDLSRALRYVDKNIMTKWGSAAYAFED HGIAOALDHPHESSIVKYPSDLPYMGSYVHAHPQMFNVAPHPALPYPTSSSFEAAP PYMSGTGIRPNTRLPFSHMPSHLGTY"	
BASE COUNT	928 a 732 c 725 g 741 t		
ORIGIN	1 bp upstream from EcoRI site.		
Query Match	52.5%; Score 760.2; DB 9; Length 3126;		
Best Local Similarity	80.7%; Pred. No. 1.8e-219;		
Matches 917; Conservative 0; Mismatches 208; Indels 12; Gaps 2			
Oy	294 TCAGGAATTCACCTGATGACTGACGCGTGCCAAAGAAGGAAAATGGTTAAGCATGTC A 353		
Dd	150 TCAAGGAATCTCCTGATGAATGCAAGTGTGCCAAAGCGGGAAGATGTGGGACGCCA 209		
Oy	354 GACATGTTGGATGAACAATAAGCAAGCTACATGGAAGAAGCATATTCCCTCCAANT 413		
Dd	210 GACACCGTTGGATGAACACTACGCAGCATACATGAGAGAAAGCATGCCACCACCACAA C 269		
Oy	414 ATGACAACCAATGAACGAAGATTATTTGCGCACAGATCCTACGTTATGAGACACAGAC 473		
Dd	270 ATGACCAAGAAAGCAGCGCAGAGATTATCGTCCACAGATCCTACGCTATGAGATACAGAC 329		
Oy	474 CATGACAGCGAGTGGCTGAGTGGGCAATGGAAGGATGTGGCTCCAGAGTGGACATC 533		
Dd	330 CATGTGGCGAGTGGCTGAGTGGGCGGTGGTAAGATATGTGGCTTCCAGACGTCAACATC 389		
Oy	534 TTGTTGTTCCAGAAATTTGATGGGAAAGATTGTGTAAATGACCAAAGATGACTTCAG 593		
Dd	390 TTGTTATTCAGAAATCGATGATGGGAAGGAACTGTGCAAGATGACCAAGAGACATTCGAC 449		
Oy	594 AGACMCAGCGGAGCTTAACGCGAGATATCTCTCGTCAACCTACACTACCTCAGAGAG 653		
Dd	450 AGGCTCACCCGAGGTATACAGCGGACATCTTCTCTCACATCTCCACTACCTCAGAGAG 509		
Oy	654 A-----GAGAGGCCACTTTATTTTTTCCAAATACATCAGTTTACCAGGAAGCAAG 704		
Dd	510 ACTCCTCTTCACATTTGACTTCAGATGATGTGTGATTAAGGCTTACAAAACTCCACAGG 569		
Oy	705 CAAGAATACACAAAGGCCAGATTACCTTATGAGCAAGGAGAGATCAGCGTGGACG 764		
Dd	570 TTAATGATGATGTAAAAACACAGATTTTACCAATATAGCCCCCAGAGATCGCTCGGACC 629		
Oy	765 AGTACAGGCCATGCCA---CTCAGTCCAAAACTTACCACCAACCATCATGCTTCAACAGTGCCC 821		
Dd	630 GTTACAGGCCACCCACAGGCCCACTTCAGAAAGCTGCTCAACATCTCTCTCCACAGTGGCC 689		
Oy	822 AAAACAGAGAAGCAGCGCTCTCACTTAGATCTTATCAGATTCTTGAAGCAGCAGCAGC 881		
Dd	690 AAAACTGAAGACCGCGTCTCACTTAGATCTTATCAGATTCTTGAAGCACAAGATGAC 749		
Oy	882 CGCTTTCAAATCCAGGAGAGTGGGCAATACAGTATGAGCGATGCTTCATGAGGACCTCTG 941		
Dd	750 CGCCTTGAATATCCAGGAGAGTGGCAATACAGCTTGGGAGCTTCTCTGAGACTCTCTG 809		
Oy	942 TCGAGACAGCTTCCAACTCCAACTGCAATCCTGGGAGGCGACAAATGGGGAAGTTCAAGATG 1001		
Dd	810 TCGGACAGCTCCAACTCCAGCTGCATCAGCTGGGAAAGGCAACCAAGGAGTTCAAGATG 869		
Oy	1002 ACAGACCTCTGATGAAGTGGCTCGGCGGTTGGGGAAGAGAAAAGCAAACTACATGAAAC 1061		
Dd	870 ACCGATCTCCGACGAGGTGGCCCGCGGTGGGGAAGGGAAGACCAACCAACCACTACATGAC 929		

Db 933 GCGTCATCACTGCGGGAAGAACCAACGAGAGTTCAAGATGACGAGACCCAGAGGCTG 992
 Oy 1020 GCTGGGGGTGGGGAGAGGAAAGCAACCTACATGACATGACAACTGACGCGT 1079
 Db 993 GCACGGGCTGGGGCGAAGCAACCAACCACTGACATGACAAAGCTGAGCGGA 1052
 Oy 1080 GCACCTGCTACTACTATGACAAATAATTTATGACTAAAGTTCTGTTAAAGCTATGCC 1139
 Db 1053 GCCCTGGCTACTACTGACAAAGACATGACAAAGTGCAGGCAAGCGGTAGCGC 1112
 Oy 1140 TACAAATTTATTTCCAGAGAAATCGTCAGGCGCTCCAGGCTCAACCTCCAGAAATCATCC 1199
 Db 1113 TACAAATTTATTTCCAGAGAAATCGTCAGGCGCTCCAGGCTCAACCTCCAGAAATCATCC 1172
 Oy 1200 ATGACAAATATCCATGACAGCTCCCTGACATGATGCTCCACATGACAGCCCAAG 1259
 Db 1173 ATGACAAATATCCATGACAGCTCCCTGACATGATGCTCCACATGACAGCCCAAG 1232
 Oy 1260 ATGAACTTTGATGCTCCCAATCCCTGCTTTGGCCGCTACATCTCAGCTTTTGTCT 1319
 Db 1233 GTCACCTGCTGCTCTCTCTCAACCTCCGTCATGCTCCGCTCACTCTCTCACTTTGCT 1292
 Oy 1320 GCCCTATATCTATCTGAAATTCACAGCTGAGGAGCATGCCCAATACAGGCTGCCA 1379
 Db 1293 CCCACAAACCCCATCTGAGCTCCCAACAGAGGATTTTACCCCAATCCAGGCTGCCA 1352
 Oy 1380 -----GCTGCTATGCTCCCTGCTCCATCTTGGGCACTGACTATA 1418
 Db 1353 CGCCATGCCAATTCACAGCTGCTCTCAACCTTGGTACTGACTATA 1400

RESULT 14
 LOCUS MEF1L 1729 bp mRNA linear ROD 02-SEP-1991
 DEFINITION Mouse Flt-1 mRNA for retroviral integration site.
 X59421
 VERSION X59421.1 GI:50974
 KEYWORDS c-ets-1 gene; ets gene family; flt-1 gene; integration site.
 SOURCE Mus musculus.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 Letwin, K.
 REFERENCE 1 (bases 1 to 1729)
 TITLE Direct Submission
 JOURNAL Submitted (08-MAY-1991) K. Letwin, Samuel Lunenfeld Res Inst at
 Mount Sinai Hospital, Div of Mol and Developmental Biol, 600
 University Avenue, Toronto Ontario M5G 1X5, CANADA
 2 (bases 1 to 1729)
 REFERENCE Ben-David, Y., Giddens, E.B., Letwin, K. and Bernstein, A.
 Erythroid leukemia induction by Friend murine leukemia virus:
 Insertional activation of a new member of the ets gene family,
 Flt-1, closely linked to c-ets-1
 JOURNAL Genes Dev. 5 (6), 908-918 (1991)
 MEDLINE 91257578
 PUBMED 2044959

FEATURES
 source
 1..1729 Location/Qualifiers
 /organism="Mus musculus"
 /strain="palb/c"
 /db_xref="taxon:10090"
 /cell_line="CB7"
 /cell_type="Friend erythroleukemia"
 /germline
 1..1729
 /gene="Flt-1"
 1..1729
 /gene="Flt-1"
 /evidence=experimental
 245..1603
 /gene="Flt-1"
 /codon_start=1
 /product="retrovirus integration site"

gene
 mRNA
 CDS

/protein_id="CAA42055.1"
 /db_xref="GI:50975"
 /db_xref="MGI:95554"
 /db_xref="SWISS-PROT:P26322"
 /translation="MDSTIEALSVSDSDSLSDSAVGAALHPKADMTASGPDYQ
 PHKINLPPOEMINOPVRNVRKREYDNRKSRSPDSCVSKCNKLVGGENAPMY
 NSYMDENKGPPTMTNERNRIVIPAPPTLMTEBHEWQMLEMALKEGLEIDTSFQ
 NMOKELCKNNKEDFLRATSAVNEVLSLSYLRSSLLAYNTSHDSSRLNVE
 DPSVDYSRGAANNNNNGLNKSPILGSGDTMKRPNRORPOPDYILGPRSSRLAP
 GSGOIQLMPLLELSDSANASCTIWTGTEGRMTDPDEVARMGRKSPNNYD
 LSRALRYDKNMTKYHGRVAYKFEFHEIALALQHPETSMYKYPDSI STYPSVH
 AHQOKVNFVPSHPSMPTSSSFEFAASQYTSPTAIYVNPSPVRHPNTHVPSHLGS
 YV"

BASE COUNT 466 a 484 c 436 g 343 t
 ORIGIN
 Query Match 39.4%; Score 570.2; DB 10; Length 1729;
 Best Local Similarity 66.8%; Pred. No. 1.1e-161;
 Matches 917; Conservative 0; Mismatches 423; Indels 33; Gaps 6;

Oy 63 ATGCAAGCACTATTAGGAAGCATTTCACTGCTGTGATGAAGCAAGCTCTGTTGAG 122
 Db 245 ATGAGCGGAGCTATTAAAGAGGCTCTGCTGTGTGATGACGATGCTCCCTTTTGAT 304
 Oy 123 TGTGCTACGAGATCG-----CCCACTTGAACAAGACAATAATGACAGCTCTCCAGT 179
 Db 305 TCAGCATACGAGAGGAGCGGACCATCTCCCAAGCAATATGACTGTTGGGGAAGTCT 364
 Oy 180 GAATATGCGCAAAACATCAAGATGAGCGCGCGCTTCCCAAGCAGAGCTGTTATACAG 239
 Db 365 GACTACGGGAGCGCCACAAAATCAACCCCTGCCACGAGCAGAGATGATCAACAG 424
 Oy 240 CCCCCGCGAGAGTTACCATTTAAGATGAGATGATACCAACCAAGTTATGGTCAAG 299
 Db 425 CC-----AGGAGAGTCAATGATCAACGGGAGATGATGACATGAATGATTCAGG 475
 Oy 300 AATTACCTGATGACTGACAGCTGGGCAAAAGAGAGAAATGTTAGCATGTTGAGCAAT 359
 Db 476 GAGTCTCGGTGAGTCACTGATGTCAGCAATGTACAAAGCTGTGGCGGAGGGAAGCC 535
 Oy 360 GTTGGGATGACTATGAGAGCTACATGAGAGAGAGCA---TATTCGCTCCAAATATG 416
 Db 536 AACCCTATGAATATATATAGCTACATGATGAGAGAGAGAGGCCCTCTCCCAACATG 595
 Oy 417 ACAACCAATAGAGAGATTTATGTCAGCAGATCTCTACGTTATGAGACAGACCAT 476
 Db 596 ACCACCAACGAGAGAGATGATGCTCTGCGAGACCCACACTGTGAGACACAGAGCAC 655
 Oy 477 GTACGGGAGTGGGTGGAGTGGGCGAGTGAAGAGATGCTTCCAGAGCTGGACATTTG 536
 Db 656 GTTCGACAGTGGCTGGAGTGGGCTATTAAGGAATACGGAATTGATGAGATTGACACTTC 715
 Oy 537 TTGTTCCAGAACTTGTGAGGAAAGAGTTGTATAATGACCAAGATGATCCACAGA 596
 Db 716 TTCTTCCAGAACTGATGAGTGGCAAGGAATGTGTAATGAAACAAGAGGAGCTCTCCGA 775
 Oy 597 CTACAGCGGAGCTATTAAGCAGATATCTCTGTACACCTACACTACTCAGAGAGAGA 656
 Db 776 GCCACCTCCGCTTACAAACAGAGAGTGTCTTGTGCGACCTCAGTTACAGGGAAG- 834
 Oy 657 GAGCGCACTTTATTTTCCAAATATCATGATGCTTACCAAGAGCAAGCAAGATATACA 716
 Db 835 -----TTCACTGCTGGCTTATTAACACACACTCCCATACAGACAGTCTTACAGACTGAAT 889
 Oy 717 ACAAGCGCATTTACTTTATGAGCAAGCAGAGATCAGCTGAGAGACTCA---GC 773
 Db 890 GTCAGGAAGACCTTCTTATGACTGTGACAGAGAGGAGCAGTGAATATATATGAAAC 949
 Oy 774 CATCCCACTAGTCAAAAGCTACCCAACTATCTTCAACAGTGGCCAAACAGAGAGAC 833
 Db 950 TCTGGCTCAACAAAGCTCTCTCTTGGAGGATCACAGACCATGAGGCAAGAACTAGAG 1009
 Oy 834 CAGCGTCTCAGTTAGATCTTATCAGATTTCTTGACCGACGACGAGCGCTTGCAAAT 893

Db	1010	CACGCGCCCGACAGATCCCTTATGACATCTGGGGCCACACGACCCCTAGCAAC	1069	QY	63	ATGGCAAGCACTATTAGAGCAATTATCAGTGTGATGAGACCAAGCTCTTTTGGAG	132
QY	894	CCAGGAGTGGGAGATACAGATATGCGAGTTCTCTACTGAGAGCTTCTGGGACAGCTCC	953	Db	121	ATGAGAGGAGCTATTAAAGAGGCTCTGTGCTGTGAGGACAGCAGCTCTCTTTGAC	180
Db	1070	CTTGGAGTGGGAGATACAGCTGTGTGAGTTCTCTGAGACTGCTCGAGAGGCC	1129	QY	123	TGTGCTACGAGATG---CCCACTTTGCAAAAGCAGAAATGACAGCTCTCTTCCAGT	179
QY	954	AACTCCAACTGCATCACCTGGGAGGACAAATGGGGAGTTCAAGATGACAGACCTGAT	1013	Db	181	TCAGCGTACGAGCAGGACCCATCTCCCAAGGCGCAGTCAGTCTGTGGGAGTCTT	240
Db	1130	AACGCGAGTGTATACCTGGGAGGAGACACGGGGAGTTCAAAATGAGGAGACCTGAT	1189	QY	180	GAATATGAGCAAAATCAAGATGAGCCGCGGTTCCCGACAGAGACTGGTATACAG	239
QY	1014	GAAATGGCTGGGCTTGGGAGAGAGAAAAGCAAACTTAACTGAACTTACAAATC	1073	Db	241	GACTACGGGAGCCCGACAGATCAACCCCTCCACACAGAGAGAGTGTATCAATCAG	300
Db	1190	GAGCTGGCCAGGGCTGGGAGAGCCGGAAGCAAGCCAAATGAAATTTATGACAACTG	1249	QY	240	CCCCCGGCGAGTTACATTAAGATGAGAGTGAACCCAAACAGGTTAATGGTCAAG	299
QY	1074	AGCGGTGACTGTGCTACTACTATGACAAATATTTATGACTAAAGTTAGTGAACGC	1133	Db	301	CC-----AGTAGAGGTCACGTCACGCGGAGATATGACATGAAATGATCCAG	351
Db	1250	AGCGGGGCGCTCCGATATCTATGACAAAACATTTATGACAAAGTGCATGGCAAGAG	1309	QY	300	AATTCACTGATATACCTACCGTGGCAAAAGAGGAAATGGTTAGCAATGACAAAT	359
QY	1134	TATGCTTACAAATTTGATTTCAAGGAAATGCTCAGGCCCTCCAGCCTCCAGAA	1193	Db	352	GAGTCTCGGTGAGCTGACAGCTTATGACAAATGAGCAAGCTGGTGGGCGGAGCTC	411
Db	1310	TATGCTTACAAATTTGATTTCAAGGAAATGCTCAGGCCCTCCAGCCTCCAGAA	1193	QY	360	GTTGGGATGAACTATGGAAGCTACATGGAAGAGAGCAT---ATTCCGCTCCAAATATG	416
QY	1194	TCATCCATGTACAAATACCATCAGACCTCCCTACATGAGTTCTTACCTGACACCCC	1253	Db	412	AACCCATGAACTACAAAGCTATATGAGAGAGAAATATGCCCCCTCTCTCCAACTG	471
Db	1370	ACATCCATGTACAAATACCATCAGACCTCCCTACATGAGTTCTTACCTGACACCAA	1429	QY	417	ACAACCAATGAAAGAGATTTATGTCAGCAGATCTTACGTTATGAGACACAGACAT	476
QY	1254	CAGAAGATGAACTTTGATGCTCCACCTCCCTGCTTGGCCGTAACCTATCAGGTTT	1313	Db	472	ACCACCAAGAGAGAGAGTATGTCCTCCGACAGCCCACTGCTGAGACAGAGCAT	531
Db	1430	CAGAAGTGAACCTTTGCTCCGCTCTCAGCCCTCTCCATGCTGCTGCTCTCCAGCTTC	1489	QY	477	GTAACGAGTGGTGAAGTGGGAGTGAAGAGATATATGTTCTTCCACAGCGACATCTG	536
QY	1314	TTTGTCTCCCTTAATCCATACGTAAGTTCACCACTGGAGGACATCAACCCAAATACG	1373	Db	532	GTAAGCAATGCTGAGAGTGGGAGTGAAGAGATATATGTTCTTCCACAGCGACATCTC	591
Db	1490	TTTGTGAGCAACATACCAATATCTGAGCTCCCACTGCTGGATCTTCCAAACCCAGT	1549	QY	537	TTTGTCCAGAACTTTATGGAAGAGTGTGTAATGACCAAAATGATCTCCAGAGA	596
QY	1374	CTGCGACGCTGCTATA-----TGCTTCCCATCTTGGCACTACTACTA	1417	Db	592	TTTTTCCAGAACTGATGAGTGGCAAGAGATGTGTAATGACCAAGAGAGACTCTCCG	651
Db	1550	GTCGCCCGCCATCTTAACACCCACGCTGCTTACACTTAGGACAGACTACTA	1602	QY	597	CTCAGCGCAGCTATTAACGAGATATCTCTCTCTCACACCTTACACTCTACAGAGAGA	656
RESULT 15	A36461	2916 bp	DNA	linear	PAT 05-MAR-1997		
LOCUS	A36461	Sequence 2 from Patent W09323549.					
DEFINITION	A36461						
VERSION	A36461.1	GI:2293779					
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
COMMENT							
FEATURES							
BASE COUNT	844 a	686 c	663 g	723 t			
ORIGIN							
Query Match	39.0%;	Score 563.8;	DB 6;	Length 2916;			
Best Local Similarity	66.5%;	Pred. No. 1e-159;					
Matches 913;	Conservative 0;	Mismatches 427;	Indels 33;	Gaps 6;			


```
Db 1186 TATGCTTACAAATTTGACTTCCAGGGCATTCGCCAGGCTTCGACGCCACATCCGACCGAG 1245
OY 1194 TCATCCATGTACAAATACCCATCCAGACCTCCCTACATGAGTTCCTACCATGCAACCCC 1253
Db 1246 TCGTCCAGTACAAAGTACCTTCTGACATCTCTACATGCGCTTCTACATGCCCCACG 1305
OY 1254 CAGAAGATGACCTTGTAGCTCCGCCATCCCGCTTGGCCGTAACTCATCCAGCTTT 1313
Db 1306 CAGAAGGTGAACTTGTCCCTCCCATCTCATCTCATGCTGTCTCTCTCTCTCTCTCTCT 1365
OY 1314 TTTGCTGCCCCCTAAATCCATCTGGAATTCACCAATGAGGATCTACCCCATATCCA-- 1371
Db 1366 TTTGAGGCCGATCATCAATATGAGACCTCCGCCAAGGGGGAATCTAACCACACCCAC 1425
OY 1372 -----GGTCCAGCTGCTCATATGCTTCCCATCTTGGCACCTACTACTA 1417
Db 1426 GTCCCGCGCATCTCTAACACCCAGGTGCTTCCACCTTAGCGCAGCTACTACTA 1478
```

Search completed: July 28, 2003, 11:20:46
Job time : 4237.82 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:39:19 ; Search time 32.1776 Seconds

(Without alignments)
9948.861 Million cell updates/sec

Title: US-09-902-772-1_COPY_650_660

Perfect score: 11

Sequence: 1 agagagagagag 11

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_un: *
28: em_vi: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rnd: *
36: em_htg_man: *
37: em_htg_vtl: *
38: em_sy: *
39: em_htg_hum: *
40: em_htg_mus: *
41: em_htg_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	20	6	AX326907 Sequence
2	11	100.0	21	6	AX244546 Sequence
3	11	100.0	21	6	AX404317 Sequence
4	11	100.0	21	6	AX404318 Sequence
5	11	100.0	22	4	DOGEP34701
6	11	100.0	22	6	AB6933 Sequence 24
7	11	100.0	22	6	AR048345 Sequence
8	11	100.0	22	6	AR079236 Sequence
9	11	100.0	23	6	AR118353 Sequence
10	11	100.0	24	10	MM884
11	11	100.0	25	6	AX115095 Sequence
12	11	100.0	32	6	AR060668 Sequence
13	11	100.0	32	6	AR102813 Sequence
14	11	100.0	36	6	AR084537 Sequence
15	11	100.0	36	6	AX113835 Sequence
16	11	100.0	42	6	A06938 F.domesticu
17	11	100.0	43	6	AR011922 Sequence
18	11	100.0	43	6	AR017793 Sequence
19	11	100.0	43	6	AR077198 Sequence
20	11	100.0	43	6	E38119
21	11	100.0	51	5	ECY17989
22	11	100.0	63	6	AR088300 Sequence
23	11	100.0	63	6	AR140197 Sequence
24	11	100.0	63	6	AR198558 Sequence
25	11	100.0	63	6	BD08437 F.fidrioblas
26	11	100.0	85	5	MPY17987
27	11	100.0	89	4	MVB2LF
28	11	100.0	90	6	AR037143 Sequence
29	11	100.0	90	6	AX435642 Sequence
30	11	100.0	90	6	165754
31	11	100.0	99	10	MUSCERBI
32	11	100.0	104	6	AR037140 Sequence
33	11	100.0	104	6	165751
34	11	100.0	105	4	MVB2LB
35	11	100.0	105	9	HUMHELAH
36	11	100.0	108	3	AY009753 Human micro
37	11	100.0	111	6	AR020453 Sequence
38	11	100.0	111	6	AR020454 Sequence
39	11	100.0	111	6	AR18349 Sequence
40	11	100.0	111	6	AR118350 Sequence
41	11	100.0	111	6	AR122315 Sequence
42	11	100.0	111	6	AR122316 Sequence
43	11	100.0	111	6	180021
44	11	100.0	111	6	180022
45	11	100.0	112	4	AF276820

ALIGNMENTS

RESULT 1
AX326907
LOCUS AX326907 20 bp DNA
DEFINITION Sequence 103 from Patent WO0178894.
ACCESSION AX326907
VERSION AX326907.1 GI:18097618
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Keith, T.
TITLE Novel human gene relating to respiratory diseases, obesity, and
inflammatory bowel disease
JOURNAL Patent: WO 0178894-A 103<25-OCT-2001>

Pred. No. is the number of results predicted by chance to have a

Genome Therapeutics Corp. (US)
Location/Qualifiers
1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer"

BASE COUNT 6 a 3 c 9 g 2 t

Query Match 100.0%; Score 11; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 4e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
2 AGAGAGAGAG 12

RESULT 2
AX244546/c 21 bp DNA linear PAT 28-SEP-2001
LOCUS Sequence 12 from Patent WO0166793.
DEFINITION AX244546
ACCESSION AX244546
VERSION AX244546.1 GI:15859473
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera.
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Vitaceae; Vitis
1 (bases 1 to 21)
Baiges Blanco,I.M.; Schaeffner,A.R. and Mas,A.
Nucleic acid and oligonucleotides derived therefrom for
specifically amplifying and specifically detecting aquaporin genes
from Vitis spec
Patent: WO 0166793-A 12/13-SEP-2003;
GSF-Forschungszentrum fuer Umwelt und Gesundheit GmbH (DE)
Location/Qualifiers
1. .21
/organism="Vitis vinifera"
/db_xref="taxon:29760"

BASE COUNT 2 a 9 c 1 g 9 t

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
20 AGAGAGAGAG 10

RESULT 3
AX404317 21 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 143 from Patent WO0224747.
DEFINITION AX404317
ACCESSION AX404317
VERSION AX404317.1 GI:21437598
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequences.
1
REFERENCE
AUTHORS Brinkmann,U. and Hoffmeyer,S.
TITLE Polymorphisms in human genes of cardiovascular regulators and their
use in diagnostic and therapeutic applications
JOURNAL Patent: WO 0224747-A 143-28-MAR-2002;
Epidaurus Biotechnologie AG (DE)
Location/Qualifiers
1. .21
/organism="synthetic construct"
/db_xref="taxon:32630"

/note="artificial sequence"
BASE COUNT 7 a 4 c 10 g 0 t

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
11 AGAGAGAGAG 21

RESULT 4
AX404318 21 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 144 from Patent WO0224747.
DEFINITION AX404318
ACCESSION AX404318
VERSION AX404318.1 GI:21437599
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequences.
1
REFERENCE
AUTHORS Brinkmann,U. and Hoffmeyer,S.
TITLE Polymorphisms in human genes of cardiovascular regulators and their
use in diagnostic and therapeutic applications
JOURNAL Patent: WO 0224747-A 144 28-MAR-2002;
Epidaurus Biotechnologie AG (DE)
Location/Qualifiers
1. .21
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="artificial sequence"

BASE COUNT 0 a 10 c 4 g 7 t

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||
11 AGAGAGAGAG 1

RESULT 5
DOG34701/c 22 bp DNA linear MAM 28-FEB-1996
LOCUS Dog (Clone: CXX-347) primer for STS 347, 5' end.
DEFINITION L24232
ACCESSION L24232
VERSION L24232.1 GI:401891
KEYWORDS
PCR identification; PCR primer; STS.
1 of 2
Canis familiaris (library: E. Ostrander, in pBluescript+) adult
Spleen DNA.
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 22)
Ostrander,E.A., Maza,F.A., Yee,M. and Rine,J.
One hundred and one new simple sequence repeat-based markers for
the canine genome
Mamm. Genome 6 (3), 192-195 (1995)
MEDLINE 95268214
PUBMED 7749226
COMMENT
Submitted by:
Fred Hutchinson Cancer Research Center
Transplantation Biology Dept
1124 Columbia; Mailstop M318
Seattle, WA 98104, USA
e-mail: EOstrander@bl.gov
PCR Buffer: PCR buffer (Perkin-Elmer/Cetus)

PCR Profile: Denaturation: 94 degrees C for 1.00 minute
Annealing: 55 or 59 degrees C for 0.45 minutes
Polymerization: 74 degrees C for 1.00 minutes
PCR Cycles: 33
Final Extension: 74 degrees C for 5.00 minutes.
Location/Qualifiers
1. .22
/organism="Canis familiaris"
/db_xref="taxon:9615"
/tissue_type="spleen"
/dev_stage="adult"
/tissue_lib="E. Ostrander, in pbluescript+"
primer_bind 1. .22
BASE COUNT 0 a 10 c 3 g 9 t
ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
DB 19 AGAGAGAGGAG 9

RESULT 6
A86933/c
LOCUS A86933 22 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 24 from Patent WO9838306.
ACCESSION A86933
VERSION A86933.1 GI:6735717
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 22)
AUTHORS Dolganov,G.
TITLE TRANSCRIPTS ENCODING IMMUNOMODULATORY POLYPEPTIDES
JOURNAL PATENT: WO 9838306-A 24 03-SEP-1998;
GENELABS TECH INC (US)
FEATURES
source Location/Qualifiers
1. .22
/organism="unidentified"
/isolate="PRIMER EGR11-6"
/db_xref="taxon:32644"
BASE COUNT 2 a 12 c 0 g 8 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
DB 15 AGAGAGAGGAG 5

RESULT 7
AR048345/c
LOCUS AR048345 22 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 24 from patent US 5821091.
ACCESSION AR048345
VERSION AR048345.1 GI:5970688
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Dolganov,G.
TITLE Method of identifying activated T-cells
JOURNAL Patent: US 5821091-A 24 13-OCT-1998;
FEATURES
source Location/Qualifiers
1. .22

BASE COUNT 2 a 12 c 0 g 8 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
DB 15 AGAGAGAGGAG 5

RESULT 8
AR079236/c
LOCUS AR079236 22 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 24 from patent US 5965427.
ACCESSION AR079236
VERSION AR079236.1 GI:10005982
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Dolganov,G. and Novikov,A.
TITLE Human RAD50 gene and methods of use thereof
JOURNAL Patent: US 5965427-A 24 12-OCT-1999;
FEATURES
source Location/Qualifiers
1. .22
/organism="unknown"
BASE COUNT 2 a 12 c 0 g 8 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
DB 15 AGAGAGAGGAG 5

RESULT 9
AR118353
LOCUS AR118353 23 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6140493.
ACCESSION AR118353
VERSION AR118353.1 GI:14099259
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 23)
AUTHORS Dower,W.J., Barrett,R.W., Gallop,M.A. and Needels,M.C.
TITLE Method of synthesizing diverse collections of oligomers
JOURNAL Patent: US 6140493-A 5 31-OCT-2000;
FEATURES
source Location/Qualifiers
1. .23
/organism="unknown"
BASE COUNT 11 a 0 c 12 g 0 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGGAG 11
DB 10 AGAGAGAGGAG 20

RESULT 10
MMB84

LOCUS NM84 24 bp DNA linear ROD 07-MAR-1997
DEFINITION M.musculus immunoglobulin heavy chain CDR3 region (germlinal centre
B8 DNA #4).
ACCESSION X67390
VERSION X67390.1 GI:50118
KEYWORDS diversity region; immunoglobulin; immunoglobulin heavy chain;
joining region; variable region.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 24)
REFERENCE 1 Jacob, J.
AUTHORS Jacob, J.
JOURNAL Submitted (14-JUL-1992) J. Jacob, University of Maryland School of
Medicine, Dept of Microbiology & Immunology, 655 W Baltimore St,
Baltimore MD 21 201, USA
2 (bases 1 to 24)
REFERENCE 2 Jacob, J. and Kelsoe, G.
AUTHORS Jacob, J. and Kelsoe, G.
JOURNAL In situ studies of the primary immune response to
MEDLINE (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for
92381435 peritonsillar lymphoid sheath-associated foci and germinal centers
1512536 J. Exp. Med. 176 (3), 679-687 (1992)
3 (bases 1 to 24)
REFERENCE 3 Jacob, J., Przybela, J., Miller, C. and Kelsoe, G.
AUTHORS Jacob, J., Przybela, J., Miller, C. and Kelsoe, G.
JOURNAL In situ studies of the primary immune response to
MEDLINE (4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region
93389394 mutation and selection in germinal center B cells
8376935 J. Exp. Med. 178 (4), 1293-1307 (1993)
COMMENT See also X67341-7, X67349-91, J00522, J00529-30, J00532-37 &
J00539.
FEATURES
source location/Qualifiers
1. .24
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/cell_line="germlinal centre B8"
/cell_type="B cells"
/tissue_type="spleen"
/dev_stage="day 10"
<1. .>24
/note="CDR3"
/codon_start=1
/protein_id="CAA47802.1"
/db_xref="GI:1333873"
/translation="ARERSSYL"
CDS
BASE COUNT 8 a 3 c 8 g 5 t
ORIGIN
Query Match 100.0%; Score 11; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11
|||||
Db 4 AGAGAGAGAG 14
RESULT 11
AX115095/c 25 bp DNA linear PAT 11-MAY-2001
LOCUS AX115095
DEFINITION Sequence 218 from Patent WO0129262.
ACCESSION AX115095
VERSION AX115095.1 GI:14032037
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 25)

AUTHORS Picoult-Newburg, J. and Pohl, M.
TITLE Genotyping reagents, kits and methods of use thereof
JOURNAL Patent: WO 0129262-A 218 26-APR-2001;
Orchid Biosciences, Inc (US)
FEATURES
source location/Qualifiers
1. .25
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer"
BASE COUNT 4 a 12 c 0 g 9 t
ORIGIN
Query Match 100.0%; Score 11; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11
|||||
Db 22 AGAGAGAGAG 12
RESULT 12
AR060668/c 32 bp DNA linear PAT 29-SEP-1999
LOCUS AR060668
DEFINITION Sequence 42 from patent US 5840839.
ACCESSION AR060668
VERSION AR060668.1 GI:5987118
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Wang, R.-F. and Rosenberg, S.A.
JOURNAL Alternative open reading frame DNA of a normal gene and a novel
TITLE human cancer antigen encoded therein
PATENT: US 5840839-A 42 24-NOV-1998;
LOCATION/Qualifiers
1. .32
/organism="unknown"
BASE COUNT 7 a 9 c 7 g 9 t
ORIGIN
Query Match 100.0%; Score 11; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGAGAGAGAG 11
|||||
Db 29 AGAGAGAGAG 19
RESULT 13
ARI02813/c 32 bp DNA linear PAT 14-FEB-2001
LOCUS ARI02813
DEFINITION Sequence 42 from patent US 6087110.
ACCESSION ARI02813
VERSION ARI02813.1 GI:12814401
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Wang, R.-F. and Rosenberg, S.A.
JOURNAL Alternative open reading frame DNA of a normal gene and a novel
TITLE human cancer antigen encoded therein
PATENT: US 6087110-A 42 11-JUL-2000;
LOCATION/Qualifiers
1. .32
/organism="unknown"
BASE COUNT 7 a 9 c 7 g 9 t
ORIGIN
Query Match 100.0%; Score 11; DB 6; Length 32;

Best Local Similarity 100.0%; Pred. No. 3.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||

Db 29 AGAGAGAGAG 19

RESULT 14

AR084537 36 bp DNA linear PAT 01-SEP-2000
LOCUS AR084537
DEFINITION Sequence 26 from patent US 5981185.
ACCESSION AR084537
VERSION AR084537.1 GI:10011308
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.

REFERENCE Unclassified.
1 (bases 1 to 36)

AUTHORS Matsun,R.S., Coassin,P.J., Rampal,J.B. and Caskey,C.Thomas.
TITLE Oligonucleotide repeat arrays
JOURNAL Patent: US 5981185-A 26 09-NOV-1999;
FEATURES Location/Qualifiers

1..36
Source /organism="unknown"

BASE COUNT 18 a 0 c 18 g 0 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||

Db 18 AGAGAGAGAG 28

RESULT 15

AX113835/c 36 bp DNA linear PAT 01-MAY-2001
LOCUS AX113835
DEFINITION Sequence 11 from Patent WO0127295.
ACCESSION AX113835
VERSION AX113835.1 GI:13940018
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 36)
artificial sequences.

AUTHORS Schandendorf,D., Paschen,A., Chakraborty,T. and Dommann,E.
TITLE Recombinant attenuated listeria for immunotherapy
JOURNAL Patent: WO 0127295-A 11 19-APR-2001;
Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
(DE)

FEATURES Location/Qualifiers

1..36
Source /organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 8 a 13 c 5 g 10 t
ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAGAGAG 11
|||||

Db 35 AGAGAGAGAG 25

Search completed: July 28, 2003, 11:20:41
Job time : 37.1776 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 07:41:54 ; Search time 429.733 Seconds
(without alignments)
7582.941 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447
Sequence: 1 gaattccgcgaacgatat.....gaagaagcgcgaagaaa 1447

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	08
Maximum Match	100%	

Database : N_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
1	1447	100.0	1447	20	AAK26551		DNA encoding chick
2	1356	93.7	1528	20	AAK26552		DNA encoding chick
3	563.8	38.9	2938	14	AAO50644		Human Hm-Fli-1 gene
4	562.2	38.9	2957	24	ABK84139		Human cDNA differ
5	544.6	37.6	2954	22	AAH02915		Human breast stress
6	387	26.7	567	22	ABA48124		Human breast cell
7	387	26.7	567	22	ABA66003		Human foetal liver
8	387	26.7	567	22	ABA33090		Human foetal liver
9	387	26.7	567	22	AAK14426		Human brain expres

C	10	387	26.7	567	22	AAK40160	Human bone marrow
C	11	387	26.7	567	22	AAI20932	Probe #10865 for g
C	12	387	26.7	567	22	AAI616176	Probe #14862 used t
C	13	387	26.7	567	22	AAI06643	Probe #6634 used t
C	14	387	26.7	567	24	ABSI44215	Human genome-deriv
C	15	354.6	24.5	473	22	ABAS401193	Human breast cell
C	16	354.6	24.5	473	22	ABAS68128	Human foetal liver
C	17	354.6	24.5	473	22	ABAS5152	Probe #13618 for g
C	18	354.6	24.5	473	22	AAK16512	Human brain expres
C	19	354.6	24.5	473	22	AAK42265	Human bone marrow
C	20	354.6	24.5	473	22	AAI48339	Probe #17025 used t
C	21	354.6	24.5	473	22	AAI086636	Probe #8687 used t
C	22	325	22.5	533	22	AAAC01342	Human secreted proc
C	23	320.4	22.1	1890	14	AAOS00662	Human Hum-flt-1-ge
C	24	276.6	19.1	549	24	ABRK41114	CDNA #54 encoding
C	25	255	17.6	420	22	ABAS45056	Human breast cell
C	26	255	17.6	420	22	ABAS45529	Human foetal liver
C	27	255	17.6	420	22	ABAS25239	Probe #3705 for ge
C	28	255	17.6	420	22	AAK03760	Human brain expres
C	29	255	17.6	420	22	AAK29224	Human bone marrow
C	30	255	17.6	420	22	AAI35180	Probe #3866 used t
C	31	255	17.6	420	22	AAI03692	Probe #3683 used t
C	32	255	17.6	454	22	ABAA42995	Human breast cell
C	33	255	17.6	454	22	ABAS3411	Human foetal liver
C	34	255	17.6	454	22	ABAS23187	Probe #1653 for ge
C	35	255	17.6	454	22	AAK01681	Human brain expres
C	36	255	17.6	454	22	AAK27129	Human bone marrow
C	37	255	17.6	454	22	AAI11718	Probe #1651 for ge
C	38	255	17.6	454	22	AAI133025	Probe #1711 used t
C	39	255	17.6	454	22	AAI016166	Probe #1637 used t
C	40	255	16.3	454	24	ABSO1685	Human genome-deriv
C	41	235.6	16.3	318	21	AACG1875	Human secreted pro
C	42	180	12.4	1752	22	AAAS13672	CDNA encoding Ratt
C	43	180	12.4	1752	24	AAI38753	Rat Lambda73 CDNA
C	44	170.6	11.8	899	23	ABLI2057	Drosophila melanog
C	45	167.8	11.6	553	22	ABAB1050	Human foetal liver

ALIGNMENTS

RESULT 1	
AA26551	
ID	AA26551 standard; DNA; 1447 BP.
XX	
AC	
XX	AA26551;
DT	
XX	14-JUN-1999 (first entry)
XX	
DE	DNA encoding chicken C-11 protein.
XX	
KW	Chicken; C-11 protein; cell calcification inhibiting activity;
KW	cell calcification inhibiting agent; c-erb protein; arthritis deformans
KW	ossification; spinal column ligament; ss.
XX	
OS	Gallus sp.
XX	
Key	Location/Qualifiers
FT	63..1418
CDS	
FT	/*tag- a
XX	
PN	JP11075871-A.
XX	
PD	23-MAR-1999.
XX	
PF	29-MAY-1998; 98JP-0166076.
XX	
PR	20-JUN-1997; 97US-005028.1
PR	18-JUN-1997; 97US-0678177.
XX	
PA	(CHUS) CHUGAI PHARM CO LTD.
XX	(UYPE-) UNIV PENNSYLVANIA.

DR WPI; 1999-257708/22
DR P-PSDB; AAY01520.

DR P-PSDB; AAY01520

PT An active protein for inhibiting cell calcification - useful for
PT measuring the calcification of a cell, for diagnosing arthritis
PT deformans or ossification of spinal column ligament

PS Disclosure; Page 7-8; 15pp; Japanese

CC The present sequence encodes a chicken C-11 protein which has cell
CC calcification inhibiting activity. The specification also describes
CC a cell calcification inhibiting agent containing c-ery protein
CC (AY01521). The proteins are used for measuring the calcification of a
CC cell, for diagnosing arthritis deformans or ossification of spinal column
CC ligament.

SQ Sequence 1447 BP; 440 A; 374 C; 317 G; 316 T; 0 other;

Query Match	100.0%	Score 1447	DB 20	Length 1447
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1447	Conservative	0	Mismatches 0	Indels 0
				Gaps 0

QY	GAATTCGGGCAAGCAATTAATTATTTATGCACTTTTGGCATCAAAATTAATTGATACA	60
Db	1 GAATTCGGGCAAGCAATTAATTATTTATGCACTTTTGGCATCAAAATTAATTGATACA	60
QY	TTATGCGCAAGCACTATTAAAGAACATTATCATGTTGAGTGAAGACCAGTCTTGTTCG	120
Db	61 TTATGCGCAAGCACTATTAAAGAACATTATCATGTTGAGTGAAGACCAGTCTTGTTCG	120
QY	AGTGTGCGTACAGGATGCGCCCGACCTTGGCAAGACGAAATGACAGCTCTCTCCAGTG	180
Db	121 AGTGTGCGTACAGGATGCGCCCGACCTTGGCAAGACGAAATGACAGCTCTCTCCAGTG	180
QY	AGTGTGCGTACAGGATGCGCCCGACCTTGGCAAGACGAAATGACAGCTCTCTCCAGTG	180
Db	121 AGTGTGCGTACAGGATGCGCCCGACCTTGGCAAGACGAAATGACAGCTCTCTCCAGTG	180
QY	AAATATGGGCAAACTCAAAAGATGAGCCCGCGGTCCCGACAGAGCTGTTATCACAGC	240
Db	181 AAATATGGGCAAACTCAAAAGATGAGCCCGCGGTCCCGACAGAGCTGTTATCACAGC	240
QY	CCCCGGCCAGAGTTACCATTTAAAGATGAGTGTAACCCAAACAGGTTAATGGGTCAAGA	300
Db	241 CCCCCGCCAGAGTTACCATTTAAAGATGAGTGTAACCCAAACAGGTTAATGGGTCAAGA	300
QY	ATTCACCTATATGCTACGACCGTGGCAAAAAGGAGAAATGTTTACGACTTCAGACATG	360
Db	301 ATTCACCTATATGCTACGACCGTGGCAAAAAGGAGAAATGTTTACGACTTCAGACATG	360
QY	TTGGAGTGAACATATGGAAGCTACATGGAAGAAGACATATTCGCGCTCCAAATGTGCAA	420
Db	361 TTGGAGTGAACATATGGAAGCTACATGGAAGAAGACATATTCGCGCTCCAAATGTGCAA	420
QY	CCAATGAACGAAGAGTTATTTGTGCCAGCAGATCTCTACGTTATGAGACAGACCATGTAC	480
Db	421 CCAATGAACGAAGAGTTATTTGTGCCAGCAGATCTCTACGTTATGAGACAGACCATGTAC	480
QY	GGCAGTGGGTGGAAGTGGCGCAGTGAAGAGATGCTCTCCAGACGTGACATCTTGTGT	540
Db	481 GGCAGTGGGTGGAAGTGGCGCAGTGAAGAGATGCTCTCCAGACGTGACATCTTGTGT	540
QY	TCCAGAACATTTATGATGGGAAAGAGTTGTTAAATGACCAAAAGATGACTTCCAGAGACTCA	600
Db	541 TCCAGAACATTTATGATGGGAAAGAGTTGTTAAATGACCAAAAGATGACTTCCAGAGACTCA	600
QY	CGCGGAGCTATTAAGCAGATATCTCTGTCACACCTTACACTACTCAGAGAGAGAG	660
Db	601 CGCGGAGCTATTAAGCAGATATCTCTGTCACACCTTACACTACTCAGAGAGAGAG	660
QY	CCACTTTTATTTTCCAAATACATGAGTTTACCCAGAAAGCAACCAAGATTAACAACAA	720
Db	661 CCACTTTTATTTTCCAAATACATGAGTTTACCCAGAAAGCAACCAAGATTAACAACAA	720
QY	GGCGAGATTATCTTATGAGCAAGGAGAGAGATCAGGCTGAGAGTCCACAGCATCTCCA	780
Db	721 GGCGAGATTATCTTATGAGCAAGGAGAGAGATCAGGCTGAGAGTCCACAGCATCTCCA	780

OY		781	CTCAGTCAAAACACTACCACCATCATCTTTCAAAGTAGTGCCCAAAAACAGAAGACACGGCTC	840
Dd		781	CTCACATCAAAACCTACCCAACCATCATCTTTCAAAGTAGTGCCCAAAAACAGAAGACACAGCGTC	840
OY		841	CTCAGATTAGATCCTTATCATGATTTCTTGGACCAACAGCAGCGCGTCTTGGAAAAATCCAGGGA	900
Dd		841	CTCAGATTAGATCCTTATCATGATTTCTTGGACCAACAGCAGCGCGTCTTGGAAAAATCCAGGGA	900
OY		901	GTGGGCAGATVACAGCTATGCGAGTTCTTACTGTGAGCTTCTGTGGACAGCTCCAACTCCA	960
Dd		901	GTGGGCAGATVACAGCTATGCGAGTTCTTACTGTGAGCTTCTGTGGACAGCTCCAACTCCA	960
OY		961	ACTGATATACCTGGGAGGGGACAATAATGGGGAATTCAAGATGACAGACCCCTGATGAAATG	1020
Dd		961	ACTGATATACCTGGGAGGGGACAATAATGGGGAATTCAAGATGACAGACCCCTGATGAAATG	1020
OY		1021	CTCGGCGTTGGGGAGAGAGAAACCAAACTTACATGATGATGATGACAAATCTCAGCCGTG	1080
Dd		1021	CTCGGCGTTGGGGAGAGAGAAACCAAACTTACATGATGATGATGACAAATCTCAGCCGTG	1080
OY		1081	CACCTTCGCTACTACTATGACAAAAATATTATGACTAAGTTCATGCTTAACGCTATGCT	1140
Dd		1081	CACCTTCGCTACTACTATGACAAAAATATTATGACTAAGTTCATGCTTAACGCTATGCT	1140
OY		1141	ACAAATTTGATTTCCAGGGAATCGCTCAGGCCCTCAGCGTCACGCTCCAGAAATGATCCA	1200
Dd		1141	ACAAATTTGATTTCCAGGGAATCGCTCAGGCCCTCAGCGCTCCAGCGTCACGCTCCAGAAATGATCCA	1200
OY		1201	TGTACAAATACCCATCAAGACCTCCCCTACATAGTTCTTACCATGACACACCCCCAGAAAGA	1260
Dd		1201	TGTACAAATACCCATCAAGACCTCCCCTACATAGTTCTTACCATGACACACCCCCAGAAAGA	1260
OY		1261	TGAATTTGTAGTCTCCCATCCCTCCCTGCTTTCGCCGTAAACCTATCCACGCTTTTTTGGCTG	1320
Dd		1261	TGAATTTGTAGTCTCCCATCCCTCCCTGCTTTCGCCGTAAACCTATCCACGCTTTTTTGGCTG	1320
OY		1321	CCCCCAATACACTAGGAATTCCACCACTGAGAGCATCACCACAAATACAGGCTGCCAG	1380
Dd		1321	CCCCCAATACACTAGGAATTCCACCACTGAGAGCATCACCACAAATACAGGCTGCCAG	1380
OY		1381	CTGCTCATATGCGCTTCCCATCTTGGCACCCTACTACTAATAGTGGGAAAGAAAGAAAGCGCC	1440
Dd		1381	CTGCTCATATGCGCTTCCCATCTTGGCACCCTACTACTAATAGTGGGAAAGAAAGAAAGCGCC	1440
OY		1441	AAGAAAA 1447	
Dd		1441	AAGAAAA 1447	
<hr/>				
RESULT 2				
ID	AXX26552	standard; DNA; 1528 BP.		
XX	AXX26552;			
DT	14-JUN-1999	(first entry)		
DE	DNA encoding chicken c-ery protein.			
RW	Chicken: C-11 protein; cell calcification inhibiting activity;			
KW	cell calcification inhibiting agent; c-ery protein; arthritis deformans;			
OS	Gallus sp.			
XX	Key	Location/Qualifiers		
FT	CDS	63..1499		
XX	JP11075871-A.	/*tag= a		
DD	23-MAR-1999.			

XX 29-MAY-1998; 98JP-0166076.
 XX 20-JUN-1997; 97US-0050297.
 XX 18-JUN-1997; 97US-0878177.
 XX (CHUS-) CHUGAI PHARM CO LTD.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX WPI; 1999-257708/22.
 XX P-PSDB; AAY01521.
 DR An active protein for inhibiting cell calcification - useful for
 XX measuring the calcification of a cell, for diagnosing arthritis
 PT deformans or ossification of spinal column ligament
 XX Disclosure; Page 8-9; 15pp; Japanese.
 XX The present sequence encodes a chicken c-ery protein. The specification
 CC also describes a chicken C-11 protein (AAY01520) which has cell
 CC calcification inhibiting activity and a cell calcification inhibiting
 CC agent containing c-ery protein. The proteins are used for measuring the
 CC calcification of a cell, for diagnosing arthritis deformans or
 CC ossification of spinal column ligament.
 XX
 SO Sequence 1528 BP; 464 A; 395 C; 330 G; 339 T; 0 other;
 Query Match 93.7%; Score 1356; DB 20; Length 1528;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1447; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 1 GAATTCGGCAAGAAATATTTATTTAGCAATTTATAGGATCAATATCTTGATACA 60
 DB 1 GAATTCGGCAAGAAATATTTATTTATTTAGCAATTTATAGGATCAATATCTTGATACA 60
 QY 61 TTATGGCAACATTTATTTAGAAAGCATTTATCATGTGTGAGTGAAGACAGCTCTTTTG 120
 DB 61 TTATGGCAACATTTATTTAGAAAGCATTTATCATGTGTGAGTGAAGACAGCTCTTTTG 120
 QY 121 AGTGTGCTTACGGATGCGCCACCTTGCAAAAGACAGAAATGACAGCTCTCTTCAGTG 180
 DB 121 AGTGTGCTTACGGATGCGCCACCTTGCAAAAGACAGAAATGACAGCTCTCTTCAGTG 180
 QY 181 AATATGGCAAAATCAAGATGATGACCCGGCGGCTTCCCAAGAGATGTTATCACAGC 240
 DB 181 AATATGGCAAAATCAAGATGATGACCCGGCGGCTTCCCAAGAGATGTTATCACAGC 240
 QY 241 CCCCGGCAAGATTACCATTTAAGTGTGATTAACCAACAGGTTAATGGGTCAGGA 300
 DB 241 CCCCGGCAAGATTACCATTTAAGTGTGATTAACCAACAGGTTAATGGGTCAGGA 300
 QY 301 ATTACCTGATGACTGACGCTGGCAAAAGAGGAAATGTTAGCAGTTCAAGCAATG 360
 DB 301 ATTACCTGATGACTGACGCTGGCAAAAGAGGAAATGTTAGCAGTTCAAGCAATG 360
 QY 361 TTGGGATGAAGTATTTGAGTGTGATGAGAAAGCATTTCCGCTCCAAATATGCAA 420
 DB 361 TTGGGATGAAGTATTTGAGTGTGATGAGAAAGCATTTCCGCTCCAAATATGCAA 420
 QY 421 CCAATGAAAGAGTATTTGTCAGCAGATCTTACGTTATGAGAGACAGACATATAC 480
 DB 421 CCAATGAAAGAGTATTTGTCAGCAGATCTTACGTTATGAGAGACAGACATATAC 480
 QY 481 GGCAGTGGCTGAGTGGGAGTGAAGAGTATGGTCTTCCAGAGTGGACATCTTGTGT 540
 DB 481 GGCAGTGGCTGAGTGGGAGTGAAGAGTATGGTCTTCCAGAGTGGACATCTTGTGT 540
 QY 541 TCCAGAACATTTGGAAGAGTGTGTAAATGACCAAGATGACTTCCAGAGACTCA 600
 DB 541 TCCAGAACATTTGGAAGAGTGTGTAAATGACCAAGATGACTTCCAGAGACTCA 600
 QY 601 CGCGGAGCTATTAAGCAGATATCTCTCTGACACCTACACTCTCAGAGAGA----- 654
 DB 601 CGCGGAGCTATTAAGCAGATATCTCTCTGACACCTACACTCTCAGAGAGACTCTC 660

DB 601 CGCGGAGCTATTAAGCAGATATCTCTCTGACACCTACACTCTCAGAGAGACTCTC 660
 QY 655 ----- 654
 DB 661 TTCACATTTGACTTCAGATGATGTGATAAGGCTTACAAACCTCTCCAGGTTAATGC 720
 QY 655 -----GAGGAGCCACTTTTATTTTCCAAATATACATGTTTACCCAGAG 699
 DB 721 ATGCTAGAAACACAGGAGAGCCACTTTTATTTTCCAAATATACATGTTTACCCAGAG 780
 QY 700 CAACGCAAAAGATTAACAAGAGCCAGATTACCTTATGAGCAAGCAGAGATCAGCGT 759
 DB 781 CAACGCAAAAGATTAACAAGAGCCAGATTACCTTATGAGCAAGCAGAGATCAGCGT 840
 QY 760 GGAAGATCACAGCATCCCACTGATCAAAAGTACCAACCATCTTCAACAGTGC 819
 DB 841 GGAAGATCACAGCATCCCACTGATCAAAAGTACCAACCATCTTCAACAGTGC 900
 QY 820 CCAAAACAGAAAGACACAGCTCTTACATGATCTTATCAGATTTCTTGACGACGCA 879
 DB 901 CCAAAACAGAAAGACACAGCTCTTACATGATCTTATCAGATTTCTTGACGACGCA 960
 QY 880 GCCGCTTGCAAAATCCAGGAGTGGGAGATACAGTATGAGAGTTCCTACTGAGAGTTC 939
 DB 961 GCCGCTTGCAAAATCCAGGAGTGGGAGATACAGTATGAGAGTTCCTACTGAGAGTTC 1020
 QY 940 TGTGGAAGCTCCAACTCCAACTGATCACTGAGGAGGACAAATGGGAGTTCAAGA 999
 DB 1021 TGTGGAAGCTCCAACTCCAACTGATCACTGAGGAGGACAAATGGGAGTTCAAGA 1080
 QY 1000 TGAAGAGCCCTGATGATGATGCTGCTGGGAGTGGGAGGAAAGAAACCTTAACATGA 1059
 DB 1081 TGAAGAGCCCTGATGATGATGCTGCTGGGAGTGGGAGGAAAGAAACCTTAACATGA 1140
 QY 1060 ACTATGCAAACTCAGCGCTGCACTTGTGCTACTATGACAAATATTTATGACTAAG 1119
 DB 1141 ACTATGCAAACTCAGCGCTGCACTTGTGCTACTATGACAAATATTTATGACTAAG 1200
 QY 1120 TTGATGTAACGCTATGCTTACAAATTTGATTTCAAGGAATGCTCAGGCTCCAGC 1179
 DB 1201 TTGATGTAACGCTATGCTTACAAATTTGATTTCCAGGAATGCTCAGGCTCCAGC 1260
 QY 1180 CTCACCCCTCAGATCATCATGATGACAAATACCATAGACCTCCCTACATGATGCT 1239
 DB 1261 CTCACCCCTCAGATCATCATGATGACAAATACCATAGACCTCCCTACATGATGCT 1320
 QY 1240 ACCATGACACCCCAAGATGAATTTGTAGCTCCCATCCCTGCTTGGCCGTAA 1299
 DB 1321 ACCATGACACCCCAAGATGAATTTGTAGCTCCCATCCCTGCTTGGCCGTAA 1380
 QY 1300 CTTTCATCAGCTTTTGTGCTGCCCCATTAATCATCTGGAATTCACCAACTGAGGACTCT 1359
 DB 1381 CTTTCATCAGCTTTTGTGCTGCCCCATTAATCATCTGGAATTCACCAACTGAGGACTCT 1440
 QY 1360 ACCCATATACAGGCTGCGCAGCTGCTCATATGCTTCCCATCTTGGACCTACTACAG 1419
 DB 1441 ACCCATATACAGGCTGCGCAGCTGCTCATATGCTTCCCATCTTGGACCTACTACAG 1500
 QY 1420 TGGGGAAGAAAGAAAGCGCCCAAGAAAA 1447
 DB 1501 TGGGGAAGAAAGAAAGCGCCCAAGAAAA 1528

RESULT 3
 AA050644
 ID AA050644 standard; cDNA; 2938 BP.
 XX AA050644;
 AC 26-MAY-1994 (first entry)
 XX Human Hum-Fil-1 gene clone BM025.
 DE

KW chromosomal translocation: chimeric; chimeraic; Ewing sarcoma;
 KW EWS gene; malignant melanoma; hum-fl-1;
 KW primitive peripheral neuroectodermal tumour; human chromosome 11;
 KW human chromosome 22; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 143..1501
 FT /*tag= a
 FT /product= HUM-FLI-1
 FT polyA-signal 2908..2913
 FT /*tag= b
 XX
 XX W09323549-A.
 XX
 XX 25-NOV-1993.
 XX
 XX 19-MAY-1993; 93WO-FR00494.
 XX
 XX 20-MAY-1992; 92FR-0006123.
 XX
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Aurias A, Delattre O, Desmaze C, Melot T, Peter M;
 XX Plougastel B, Thomas G, Zucman J;
 XX
 XX WPI: 1993-386580/48.
 XX P-PSDB; AAR44556.
 DR
 XX
 XX New nucleic acid of EWS gene and its hybrid(s) - contg. gene
 PT sequence involved in chromosomal trans-location, also derived
 PT of mRNA, probes, fusion proteins etc., for diagnosis and treatment
 PT of Ewing sarcoma and melanoma
 XX
 XX PS Disclosure; Fig 7; 123bp; French.
 XX
 XX The probe 11R1 was used to screen a human marrow cDNA library
 CC (Clontech cat. # HL1058). The clone BM025 was identified and
 CC sequenced. It represents the entire coding region together with
 CC 5'- and 3'-UTRs of the Hum-Fl-1 gene.
 XX
 SQ Sequence 2938 BP; 847 A; 692 C; 675 G; 724 T; 0 other;

Query Match 39.0%; Score 563.8; DB 14; Length 2938;
 Best Local Similarity 66.5%; Pred. No. 4.8e-175;
 Matches 913; Conservative 0; Mismatches 427; Indels 33; Gaps 6;

QY 63 ATGGCAAGCACTTTAAGAGACATATCATGTGTGATGAAGACCAAGTCTTGTAG 122
 DB 143 ATGGACGGGACTATTAAGAGGCTGTGTGTTGAGCGAGACCAAGTCTTGTAG 202
 QY 123 TGTGCTACGATG---CCCCACCTTGAAGACAGAAATGACAGCTCTTCACGT 179
 DB 203 TACGCGTAGCGAGCGGACGCCATCTCCCAAGCGGACATGACTGCTGGGAGTCT 262
 QY 180 GAATATGAGCAACATCAAGATGAGCCGCGGTTCCCGACAGAGTGGTTATACAG 239
 DB 263 GACTAGGGGAGCCCAAGATCAACCCCTCCACACAGAGAGTGATCAATACAG 322
 QY 240 CCCCCCGCAGATTTACATTAAGATGAGATGATACCCAAACAGTTAATGGTCAAG 299
 DB 323 CC-----AGTGGGGGTCAACGTCAACGGGAGATATGCCATGAATGATCCAGG 373
 QY 300 AATTCACTGATACGACGCGTGGCAAAAGAGGAAATGGTTAGCACTTACAGCAAT 359
 DB 374 GATGTCGCGTGAAGTACGACGCTTACCAATTCAGCAAGCTGTGGGCGGAGCGAGTCC 433
 QY 360 GTTGGATGAATATGAAAGCTACATGAGAGAGAGCAT--ATTCCGCTCCAAATATG 416
 DB 434 AACCCCATGAATCAACAAGCTATATGAGAGAGAGATGGCCCTCTCCCAACATG 493
 QY 417 ACACCAATGAACGAAGATTATTGTGCCAGCATCTTACGTTATGAGACAGACCAT 476

DB 494 ACCGACACAGAGAGAGATCATGCTCCCGCAGACCCACACACTGTGTGACAGAGACAT 553
 QY 477 GTACGGCAGTGGCGGAGTGGGACGATGAGAGATGCTCTCCAGCGTGCATCTTG 536
 DB 554 GTGAGGCAATGCTGGGATGGCCATAAGAGATATGCTTGTATGGAGATGACATCTCC 613
 QY 537 TTTTCCAGAACATTTGATGGGAAAGAGTTGTAAATGACCAAAAGTGAATTCAGAGA 596
 DB 614 TTTTCCAGAACATTTGATGGGAAAGAGTGTAAATGAAACAAAGAGACTTCTCCGCG 673
 QY 597 CTCAGCGGAGCTATTAACGAGATATCTCTCTGACACCTTACACTACCTACAGAGAGA 656
 DB 674 GCCACGACCTCTTAACACAGGAGATGCTGTGTGACACCTCACTAGTACCTCAGGAAAG- 732
 QY 657 GGAGCCACTTTTATTTTCCAAATATCATGATTTACCCAGAGACCAAGCAAGATACA 716
 DB 733 -----TTCACTGCTGGCTTAAATACAACTCCACACGACCAATCTTCACTGTGAT 787
 QY 717 ACAAGGCCAGATTTACCTTATGAGCAAGCGAGAGATCAGCTGTGACAGTCAAGCCAT 776
 DB 788 GTCAAGAGAGACCTTCTTATGACTGATGAGAAAGAGACATGGGCAATTAACATGAT 847
 QY 777 CC---CACTCAGTCAAAAGCTACCCAACTCATCTTCAACAGTCCCAAAACAGAAAG 833
 DB 848 TCTGCTCTCAACAAAGTCTCTCCCTGGAGGGGACAAACGATCACTAAGAAATACAG 907
 QY 834 CAGCGTCTCAGTTAGTACCTTATGACATCTTGAGCCGACGACGCGCTTGCAAT 893
 DB 908 CAAGGCGCCAGAGATCTCTTATGATCAATCTTGAGCCGACGACGCTCTGACAGC 967
 QY 894 CAGAGAGTGGGAGATACAGTATGAGCAGATCTCTTCTGAGAGCTTGTGTGACAGCTCC 953
 DB 968 CTGGAAGAGGGGAGATCAAGCTGTGAGCATCTCTCTGAGAGTGTCTCTCCAGCGCC 1027
 QY 954 AACTCCACTGATCACTCTGAGGGGACAAATGGGAGTTCAAGATGACAGACCTGAT 1013
 DB 1028 AAGCGACGCTGTATCACTCTGAGGAGGACCAAGGGGAGTTCAAAATGACAGGACCCGAT 1087
 QY 1014 GAAGTGGCTGGGGGTTGGGAGAGAGAAAGCAAACTTAACTGAACTGTGCAAACTC 1073
 DB 1088 GAGGTGGCCAGGGGCTGGGGGAGCGGAAAGCAAGCCCAACATGAATTCGACAAAGCTG 1147
 QY 1074 AGCCGTGACTTGTGCTACTATGACAAATAATTTATGATGAATGATCTATGATAAGCG 1133
 DB 1148 AGCCGGGCGCTCGTATATGATGAATAAACATTATGACAAAGTGCAGGCAAAAGA 1207
 QY 1134 TATGCTACAAATTTGATTTCCACGGAATGCTCAGAGCCCTCCAGCTCACCTCCAGAA 1193
 DB 1208 TATGCTTACAAATTTGATTTCCACGGAATGCTCAGAGCCCTCCAGCTCACCTCCAG 1267
 QY 1194 TCATCATGTACAAATTAACCATCAGACCTCCCTACATGATGTTCTTACATGACACCCC 1253
 DB 1268 TCGTTCATGTACAAAGTACCTTCTGTGATCTCTTACATGCTTCTTACATGCTCCAC 1327
 QY 1254 CAGAAGATGAATTTGATGCTCCCATCCCTGCTTGGCCGTAACTTCATCACTT 1313
 DB 1328 CAGAAGTGAATTTGCTCCCTCCCATTCATCTCCATGCTCTTCACTTCTCCAGCTT 1387
 QY 1314 TTTGCTGCCCTTATTCATCTATGAAATTACCAACTGAGAGCTCTTACCCCAATACCA-- 1371
 DB 1388 TTTGAGGCGGATCAAAATACTGAGCTCCCTCCCAAGGGGGAATTCACCCCAACCCCAAC 1447
 QY 1372 -----GGTGGCAGCTGTGCTATGATGCTTCCATCTTCCAGCTTACTACTA 1417
 DB 1448 GTCCCGGCGCATCTTAAACACCAAGTCCCTTACACTTGAAGCAGTACTACTA 1500

RESULT 4
 ABR84139
 ID ABR84139 standard; cDNA; 2957 BP.
 XX
 AC ABR84139;

XX 14-AUG-2002 (first entry)

DE Human cDNA differentially expressed in granulocytic cells #710.

XX

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;

XX viral infection; parasitic infection; protozoal infection;

XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

XX cardiac reperfusion injury; renal reperfusion injury; AIDS;

XX adult respiratory distress syndrome; inflammatory bowel disease;

XX Crohn's disease; ulcerative colitis; periodontal disease;

XX granulocyte activation; chronic inflammation; allergy.

XX

OS Homo sapiens.

XX MO20228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US30821.

XX 03-OCT-2000; 2000US-237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX MPI; 2002-435328/46.

XX

PT Detecting granulocyte activation by detecting differential expression

PT of genes associated with granulocyte activation, which serves as

PT diagnostic markers that is useful for monitoring disease states and

PT drug toxicity.

XX

XX Claim 1; SEQ ID NO 710; 114pp; English.

XX

XX The invention relates to detecting (M1) granulocyte (GC) activation

XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by

XX DNA chip analysis as given in the specification, and comparing

XX the expression level to an expression level in an unactivated

XX GC, where differential expression of Gs is indicative of GCA.

XX Also included are modulating (M2) GA by contacting GC with an agent

XX that alters the expression of at least one gene in Gs; (2) screening (M3)

XX for an agent capable of modulating GCA or an inflammation (especially

XX chronic) in a tissue, an allergic response in a subject, exposure of a

XX subject to a pathogen or sterile inflammatory disease using the

XX gene expression profile; (3) detecting (M4) an inflammation (especially

XX chronic) in a tissue, an allergic response in a subject, exposure of a

XX subject to a pathogen or sterile inflammatory disease, by detecting the

XX level of expression in a sample of the tissue of gene(s) from Gs, where

XX the level of expression of the gene is indicative of inflammation;

XX (4) treating (M5) an inflammation (especially chronic) or in a tissue,

XX an allergic response in a subject, exposure of a subject to a pathogen

XX or sterile inflammatory disease, by contacting a tissue having

XX inflammation with an agent that modulates the expression of gene(s)

XX from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for

XX modulating Gs; M3 is useful for screening an agent capable of modulating

XX GCA preferably in an inflammation in a tissue; M4 is useful for

XX detecting an inflammation (especially chronic) in a tissue, an allergic

XX response in a subject, exposure of a subject to a pathogen or sterile

XX inflammatory disease (e.g. psoriasis, rheumatoid arthritis,

XX glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal

XX reperfusion injury, AIDS, adult respiratory distress syndrome,

XX inflammatory bowel disease, Crohn's disease, ulcerative colitis,

XX periodontal disease; also bacterial infection, viral infection,

XX parasitic infection, protozoal infection, fungal infection and M5 is

XX useful for treating one of the above conditions. The present

XX sequence represents a gene differentially expressed in granulocytes.

XX Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format directly from WIPO at

XX ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 2957 BP; 855 A; 698 C; 680 G; 724 T; 0 other:

SO Query Match 38.9%; Score 562.2; DB 24; Length 2957;

Best Local Similarity 66.4%; Pred. No. 1.6e-174;

Matches 912; Conservative 0; Mismatches 428; Indels 33; Gaps 6;

QY 63 ATGCGAAGCACTATTAAAGAAAGCTTTATCACTGGTGAAGACACAGCTCTTGTGGAG 122

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 173 ATGACGGGACTATTAAAGAGGCTGTGCGTGTGAGCGGACGACGATCCCTCTTGAC 232

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 123 TGTGCTACGAGATG-----CCCCACCTTGCAAGACAGAAATGACAGCTCTTCCAGT 179

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 233 TCAGCGTACGGAGGGGAGGCCATCTCCCAAGGCCACATGATGCTCTGGGGAGTCT 292

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 180 GAATATGGGCAAAACATCAAGATGAGCCGGCTTCCCGACGAGCTGTTATCAAG 239

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 293 GACTACGGGAGCCGCCAAGATCAACCCCTCCACACGACGAGAGTGAATCAATCAG 352

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 240 CCCCCGGGCGAGATTACCATTTAAGTGAAGTGAACCAACGAGTTAAGGTCAGG 299

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 353 CC-----AGTAGAGGTCAACGTCAAGCGGAGTAGTACACATGAATGATCCAGG 403

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 300 AATTCACCTGATGACTGAGCGGTGCAAAAGAGGAGAAATGTTAGAGTTCAGACAT 359

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 404 GAGTCTCGGTGAGCTGACGCGTTAGCAAAATGCAAGAGCTGTGGGCGAGCGGAGTCC 463

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 360 GTTGGGATGAACCTATGGAAGCTACATGGAAGAGAGCAT--ATTCCGCTTCAAAATG 416

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 464 AATCCATGAACCTAACAGCTATATGAGACAGAAAGATGAGCCCTCTCCCAACATG 523

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 417 ACAACCAATGACCAAGAGTATTTGTCGACGACATCTTACGTTATGAGACAGACAT 476

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 524 ACCACCAAGAGAGAGATGATCTGTCGCCCAACCCACACATGATGACACAGAGCAT 583

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 477 GTACGGAGTGGCTGGAGTGGGAGTGAAGAGTATGCTTCCAGAGTGAATCTTG 536

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 584 GTGAGGCAATGCTGAGTGGAGTGGCCATAAAGAGTACAGCTTGATGAGATGACATCC 643

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 537 TTGTTCCAGAAATGATGAGGAAGAGTGTGTAATAATGACCAAAATGATCTTCAGAGA 596

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 644 TTTTTCAGAAATGATGAGTGAAGAGTGTGTAATAATGAAAGAGAGGACTCTCCGCG 703

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 597 CTCAGCGCGAGCTTAAACGAGATATCTCTGTGACACCTACACTACTCTACAGAGAGA 656

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 704 GCCACCACTCTTCAACACAGGAGTGTGTCACACCTCAGTACTCAGGAAAG- 762

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 657 GGAGCCACTTTATTTTCCAAATACATCACTTACCCAGAGCAAGCAAAAGATACA 716

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 763 -----TTCACTGCTGGCCTATATATACACCTCCACACCGACCAATCTCGATTGAGT 817

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 717 ACAAGGCGAGATTATCTTATGAGCAACGAGAGATCAGGTGAGAGTCAAGAGCAT 776

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 818 GTCAAGAGAGACCTCTTATGACTGTCGAGAGAGAGAGTGGGCAATATACATGAT 877

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 777 CC---CACTAGTAAAGAGTACCCCAACATCTTCAACAGTCCCAAAACGAAGAC 833

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 878 TCTGGCTCAACAAAGATCTCTCCCTTGAGGGGGCACAAAGATAGTAAGAAATACAGAG 937

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 834 CAGGCTCTCATAGTAGATCTTATCAAGATCTTGGACGAGACACAGCCGCTGCAAT 893

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 938 CACGCGCCCAAGCCAGATCCGTATCAATCTTGAGCCGACACAGATGCGCTAGCCAAC 997

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 894 CCAAGGAGTGGGAGATACAGTATGAGCACTTCTACTGAGAGCTTGTGGAGACAGCTCC 953

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 998 CTTGGAAGGGGAGATCAAGCTGTGGCAATCTCTCTGGAGTGTCTCCGACAGGCGC 1057

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 954 AATCTCACTGATCACTCTGGAGGAGCAAAATGGGAGATTCAGAGATGACAGACCTGAT 1013

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 1058 AACGCCAGCTGTATCACTCTGGAGGAGGAGCAACGCGGAGTTCAAAATGAGGAGCCGAT 1117

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 1014 GAAGTGCTCGGCTTGGGAGAGAGAAAGCAAACTTAACATGATGACAAATC 1073

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	1118	GAGGTGGCCAGGGCGCTGGGGGAGCGGAAAAAGCAAGCCCAACATGAAATTAGACAAAGCTG	1177
QY	1074	AGCCGTGCACTTCGCTCACTACTATGACAAAAAATATTATGACTAAAGTTTCATGTAAAGCG	1133
Db	1178	AGCGGGGCGCCCTCGTTATTTACTATGATAAATAAACATTATTGACCAAAAGTGACAGCGCAAGA	1237
QY	1134	TATGCGTACAAATTTGATTTCCACAGGAATGCGTAGGCGCCCGACGCTCAGCCACCCTCAGAA	1193
Db	1238	TATGCTTACAAATTTTGACTTCCACGAGGATATGGCCAGGCTCTGCAGCCACATCCGACGGAG	1297
QY	1194	TCATTCATGTACAAATTAACCATCAGACCTCCCTACATAGATTTCTTACATGACACGCC	1253
Db	1298	TGCGTCATGTACAAAGTAGTACCCCTTTCGACATCTCTACATGCGCTTCTTACATAGCCACACG	1357
QY	1254	CAGAAGATGAACCTTGTAGCTCCCCATCCCTCGCTTTGGCCCGTAACCTCAATCCAGGCTTT	1313
Db	1358	CAGAAGGTGAACCTTGTCCCTCCCATTCATCTCCATGCGCTGTCACTTCTCCAGGCTTC	1417
QY	1314	TTTTCGTGCCCCCTAATTCATACTGGAATTCACCAACTGGAGGACTCTACCCCAATACA--	1371
Db	1418	TTTTCGAGCCCGCATCACAAATACTGGACCTCCGCCAGGGGGGAATCTACCCCAACCCCAAC	1477
QY	1372	-----GGGTGCCAGCTGCTCATATGCTCTCCCATCTTGGGACACTACTACTA	1417
Db	1478	GTCCCCCGCATCTTAACACCCAGCTGCTTCACTATTGGACACTACTACTA	1530

RESULT 5
AAH02915
ID AAH02915 standard: DNA: 2954 BP

AC AAH02915;

DT 15-JUN-2001 (first entry)

Human shear stress-response coding sequence SEQ ID NO: 83.

KW Human; shear stress-response protein; vascular disease,
arteriosclerosis; ds.

OS Homo sapiens.

PN W0200125427-A1.

PD 12-APR-2001.

PF 02-OCT-2000; 2000WO-JP06840.

PR 01-OCT-1999; 99JP-0280976.

PA (KYOW) KYOWA HAKKO KOGYO KK.

[illegible]

PI Kuga T, Sekine S, Nakamura Y, Sugano S;

DR WPI; 2001-266308/27.

XX

PT useful in diagnosis and treatment of vascular disease caused by

[illegible]

claim 20; Page 402-400; 0/0pp; uapaltese

the present invention provides the protein and coding sequences of a number of human shear stress response proteins. These are useful in the

CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis including heart failure, post-DMCA stenosis and

CC hypertension.
XY

SQ Sequence 2954 BP; 846 A; 698 C; 683 G; 727 T; 0 other;

Query Match	37.68;	Score 544.6;	DB 22;	Length 2954;
Best Local Similarity	66.18;	Pred. No. 1.1e-168;		
Matches 908; Conservative	0;	Mismatches 429;	Indels 36;	Gaps 7;

QY	63	ATGGGAAACACATATTAAAGAAACATTAATCAGTGGGTGAAGAACACATCTCTGGTTAG	112
Db	173	ATGGAGGAGACATTAATAAGAGGCTGTGTGGGTGAGACGACGACCACTCCCTTTTGAC	232
QY	123	TGTGCTACGGATCG---CCCACCTTGCAAAAGACAAATAAGACAGCCTCTCTCCAGT	179
Db	233	TCAGGGTACGGGCGCGCAGGCCATCTCCCAAGGCGCATATGACTGCTCGGGAGTCT	292
QY	180	GAATATGGGCAACATCAAGATGAGCCCGCGCTTCCCAAGCAGAGCATGGTTATACAG	239
Db	293	GACTACGGGACAGCCCAAGATCAACCCCTCCACACAGCAGAGATGATCAATACG	352
QY	240	CCCCCGGCGAGTTACCATTTAAGATGAGTGTAAACCAACACAGTTAATGGGTCAAG	299
Db	353	CC-----AGTGGGGTCAACGTCAAGCGGGGAGTATGACACATCAATGAATCCAGG	403
QY	300	AATTCACCGATGACTGCGAGGTGGCAAAAGAGGAAATATGGTATGAGATTCAACAT	359
Db	404	GAGTCTCGGTGGACTGCAAGGTTTACAAATGACCAACGTGTGGTGGGAGCGCATCC	463
QY	360	GTTGGGATGAACTATGGAAGCTACATGGAGAGAGCAT---ATTCCGCTCCAAATATG	416
Db	464	AAACCCATGAACGTATATACACAGCTATATGGACGAGAAATGGGCCCCCTCTCCCAATC	523
QY	417	ACAACCATGAACGAGATTATTTGCGACAGATCTTACGTTATGGACGACAGCAT	476
Db	534	ACACCAACGAGAGAGAGATCTCTCCCGCAGACCCACATCTGTGACACAGAGCAT	583
QY	477	GTACGCGATGGCTGAGTGGGACGTGAAGGATGTGTGTTCCGACGATGACATCTTG	536
Db	584	GTGAGGCAATGGCTGAGTGGGCCATTAAGGAGTACAGCTTGATGGAGATGACACATCC	643
QY	537	TTGTTCCGAAACATTGATGGAGAAAGATTGTGTAAATGACCAAAAGATGACTTCCAGAG	596
Db	644	TTTTTCCAGAACATGATGATGGCAAGAACATGTGTAAATGAACAAAGAGATCTTCCGCG	703
QY	597	CTCAGGCGGAGCTATTAACGACAGATATCTCTGTGCACACTACACTACTCTAGAGAGAG	656
Db	704	GCCACACACCTCTACAAACGGAAGTGTGTGTACACCTCAGTTACCTTAGGGAAG-	762
QY	657	GGAGCCACTTTATTTTCCAAATACATCACTTATACCGAGAGCAACGCAAAATATAA	716
Db	763	-----TTACGTGCTGGCCTATTAATTAACAACCTCCACACCGACCAATCTCTACGATTGAGT	817
QY	717	ACAAGGCGAGATTACCTTATGAGCAAGCAGAGAGATCAGCGTGGACGAGTACACGCAT	776
Db	818	GTCAAGAAGACCCCTTCTTATGACATCACTCAGAAAGAGAGAGCGTTGGGCAATATCATGANT	877
QY	777	CC---CACTCAGTCAAAAGCTAACCCACACCATATCTTCAACAGTGGCCCAAAACGAAGAC	833
Db	878	TCTGGCCTCAACAAAAGTCTCCCTTGGAGGGGACAAACGATCAGTAAACATACAGAG	937
QY	834	CAGCGTCCCTAGTTAGATCTTTATCAGATTTTGGACCGACGACGCGCTTTGCAAT	893
Db	938	CAAGGGCCCCACGACGATCCGATACAGATCTCTGGGGCCGACACAGATGCGCTAGCAAC	997
QY	894	CCAGGAGATGGGCGAGATACAGCTATGCGAGTTTCTACTGTGAGCTTGTGTGGACAGCTCC	953
Db	998	CTGTGAACCGGCGAGATTCACAGCTGTGGCAATTTCTCTCGGTGGAGCTGTCTCCGACAGGCCC	1057
QY	954	AACTCCAACTCATACCTGGGAGGGGCAATGGGAGGTCAAGATGACAGACCCGTAT	1013
Db	1058	AAACGCAAGCTATATCCTGGGAGGGGACCAACGGGGGATTCAAAATATACGACCCGAT	1117
QY	1014	GAAGTGCCTGCGGCTTGGGAGAGAGAAAAGCAAACTTAACATGAACTATGACAAACTC	1073
Db	1118	GAGGTGGCGAGGCGCTGGGGCGACGAGAAAAGCAAGCCCAACATGAAATACGACACAGCTG	1177
QY	1074	AGCCGTGACATTGCTACTACTATGACAAAAATATATGACTAAAGTTGATGTTAAAGCC	1133

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 DR Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow -
 XX
 XX Example 4; SEQ ID NO: 14717; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
 Query Match 26.7%; Score 387; DB 22; Length 567;
 Best Local Similarity 83.7%; Pred. No. 5.2e-117;
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
 QY 896 AGGAGTGGGACATACGCTATGCGAGTCTCTACTGAGCTTCTGTGAGACGCTCAA 955
 DB 523 AGGACAGTGGGACATACGCTATGCGAGTCTCTCTGAGCTCTGTGAGACGCTCAA 464
 QY 956 CTCACACTGCATCCTCGGAGGAGGACCAATGAGGAGTTCATGACACACCTGATGA 1015
 DB 463 CTCACACTGCATCCTCGGAGGAGGACCAATGAGGAGTTCATGACACACCTGATGA 404
 QY 1016 AGTGGCTCGGCTGTGGGAGAGAGAAAGCAACCTACATGATGATGACAAACTGAG 1075
 DB 403 GGTGGCCCGGCGTGGGAGAGAGAGAAAGCAACCTACATGATGATGACAAACTGAG 344
 QY 1076 CCGTGCACCTTGGCTACTACTATGACAAATATATGATGATGATGATGATGATGAT 1135
 DB 343 CCGGCGCTTCCGCTACTACTATGACAAATATATGATGATGATGATGATGATGATGAT 284
 QY 1136 TGCCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGCTCCAGATC 1195
 DB 283 CCGCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGCTCCAGATC 224
 QY 1196 ATCCATGTAACAAATACCATCAGACCTCCCTCATGATGATGATGATGATGATGATGAT 1255
 DB 223 ATCTGTACAAATACCATCAGACCTCCCTCATGATGATGATGATGATGATGATGATGAT 164
 QY 1256 GAAGATGAATTTGTAGTCCCATCCCTGCTTGGCCGTAACTCCAGCTTTT 1315
 DB 163 GAAGATGAATTTGTAGTCCCATCCCTGCTTGGCCGTAACTCCAGCTTTT 104
 QY 1316 TGCCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGATC 1375
 DB 103 TGCCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGATC 44
 QY 1376 GCCAGCTGCTCATATGCTTCCCTGCTTGGCACTGCTACTATA 1418
 DB 43 CCCACAGCAGCATATGCTTCCCTGCTTGGCACTGCTACTATA 1
 RESULT 11
 AAI20932/c
 ID AAI20932 standard; DNA; 567 BP.
 XX
 XX AAI20932:
 AC
 XX
 XX 12-OCT-2001 (first entry)
 DE Probe #10865 for gene expression analysis in human cervical cell sample.
 XX Probe: human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX

OS Homo sapiens.
 XX
 XX W0200157278-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001W0-US006070.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 XX
 DR Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human cervical epithelial cells -
 XX
 XX Claim 25; SEQ ID NO 10865; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human HeLa cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
 Query Match 26.7%; Score 387; DB 22; Length 567;
 Best Local Similarity 83.7%; Pred. No. 5.2e-117;
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
 QY 896 AGGAGTGGGACATACGCTATGCGAGTCTCTACTGAGCTTCTGTGAGACGCTCAA 955
 DB 523 AGGACAGTGGGACATACGCTATGCGAGTCTCTCTGAGCTCTGTGAGACGCTCAA 464
 QY 956 CTCACACTGCATCCTCGGAGGAGGACCAATGAGGAGTTCATGACACACCTGATGA 1015
 DB 463 CTCACACTGCATCCTCGGAGGAGGACCAATGAGGAGTTCATGACACACCTGATGA 404
 QY 1016 AGTGGCTCGGCTGTGGGAGAGAGAAAGCAACCTACATGATGATGATGATGATGATGAT 1075
 DB 403 GGTGGCCCGGCGTGGGAGAGAGAGAAAGCAACCTACATGATGATGATGATGATGATGAT 344
 QY 1076 CCGTGCACCTTGGCTACTACTATGACAAATATATGATGATGATGATGATGATGATGAT 1135
 DB 343 CCGGCGCTTCCGCTACTACTATGACAAATATATGATGATGATGATGATGATGATGAT 284
 QY 1136 TGCCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGATC 1195
 DB 283 CCGCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGATC 224
 QY 1196 ATCCATGTAACAAATACCATCAGACCTCCCTCATGATGATGATGATGATGATGATGAT 1255
 DB 223 ATCTGTACAAATACCATCAGACCTCCCTCATGATGATGATGATGATGATGATGATGAT 164
 QY 1256 GAAGATGAATTTGTAGTCCCATCCCTGCTTGGCCGTAACTCCAGCTTTT 1315
 DB 163 GAAGATGAATTTGTAGTCCCATCCCTGCTTGGCCGTAACTCCAGCTTTT 104
 QY 1316 TGCCTCAAAATTTGATTTCCAGGAATCGCTCAGGCGCTCCAGCTCCAGATC 1375

```

Db      103 TGCATCCCCAACCATATCTGGAAATTCACCAACTGGGGTATATACCCCAACACTAGGCT 44
QY      1376 GCCAGCTGCTATATGCTTCCCATCTTGGACACTACTACTA 1418
Db      43 CCCACACGACCATATGCTCTCTCATCTGGGACACTACTACTA 1

RESULT 12
AAI46176/c
ID      AAI46176 standard; DNA; 567 BP.
AC      AAI46176;
DE      17-OCT-2001 (first entry)
XX      Probe #14862 used to measure gene expression in human placenta sample.
XX      Probe; microarray; human; placenta; antenatal diagnosis.
XX      Probe; microarray; human; placenta; antenatal diagnosis; ss.
XX      genetic disorder; ss.
XX      Homo sapiens.
XX      WO200157272-A2.
XX      09-AUG-2001.
XX      30-JAN-2001; 2001WO-US000663.
XX      04-FEB-2000; 2000US-0180312.
XX      26-MAY-2000; 2000US-0207456.
XX      30-JUN-2000; 2000US-0608408.
XX      03-AUG-2000; 2000US-0632366.
XX      21-SEP-2000; 2000US-0234687.
XX      27-SEP-2000; 2000US-0236359.
XX      04-OCT-2000; 2000GB-0024263.
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
XX      WPI; 2001-488897/53.
XX      Human genome-derived single exon nucleic acid probes useful for
XX      analyzing gene expression in human placenta.
XX      Claim 25; SEQ ID No 14862; 654pp; English.
XX      The present invention relates to single exon nucleic acid probes (SENP).
XX      The present sequence is one such probe. The probes are useful for
XX      producing a microarray for predicting, measuring and displaying gene
XX      expression in samples derived from human placenta. The probes are useful
XX      for antenatal diagnosis of human genetic disorders.
XX      Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other:
SQ
Query Match      26.7%; Score 387; DB 22; Length 567;
Best Local Similarity 83.7%; Pred. No. 5.2e-117;
Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY      896 AGGAGTGGGACAGATACGATGAGTTCCTCTGAGTTCCTGTCGACACTCCAA 955
Db      523 AGCGATGGCCAGATCCAGCTTTGGCACTTCCTCTGAGTTCCTGTCGACACTCCAA 464
QY      956 CTCCAACTGCATCACTTGGAGGAGGACCAATGGGAGTTCAAGATGACAGACCTGATGA 1015
Db      463 CTCAGCTGCATCACTTGGAGGAGGACCAAGGGAGTTCAAGATGACAGATCCGACGA 404
QY      1016 AGTGGCTGGGCTGGGAGAGAGGAAAGCAAACTTACGATGACAACTGAC 1075
Db      403 GGTGGCCCGGCTGGGAGAGAGGAGCAAAACCCAACTGACATGACGTAAGCTCAG 344
QY      1076 CCGTGCACTTGGCTACTACTATGACAAAATAATATTATGACTAACTTATGTTAAACGCTA 1135

```

```

Db      343 CCGGCCCTCCGTTACTACTATGACAGAAATATCATGACAAAGGTCATGGGAACGCTA 284
QY      1136 TGCTCAAAATTTGATTTCCACGAGATGCTGAGCCCTCCAGCCTCACTCCAGAAATC 1195
Db      283 CGCTCAAAAGTTGCACTTCCACGAGATGCTGAGCCCTCCAGCCTCACTCCAGAAATC 224
QY      1196 ATCCATGACAAATATCCATGACATCCCTCCCTTACATGATGCTTACATGACATCCCA 1255
Db      223 ATCTGTGACAAAGTACCTCCCTCCCTGACATGAGGCTCTATGACAGCCACCA 164
QY      1256 GAAGATGAATTTTGTAGCTCCCATCCCTGCTTTGGCCGTAACCTCAGCTTTT 1315
Db      163 GAAGATGAATTTTGTAGCTCCCATCCCTGCTTTGGCCGTAACCTCAGCTTTT 104
QY      1316 TGCTGCCCCATATGCTGGAATTCACCACTGAGGACATTCACCCAAATACCAGCT 1375
Db      103 TGCATCCCCAACCATATCTGGAAATTCACCAACTGGGGTATATACCCCAACACTAGGCT 44
QY      1376 GCCAGCTGCTATATGCTTCCCATCTTGGACACTACTACTA 1418
Db      43 CCCACACGACCATATGCTCTCTCATCTGGGACACTACTACTA 1

RESULT 13
AAI06643/c
ID      AAI06643 standard; DNA; 567 BP.
AC      AAI06643;
DE      09-OCT-2001 (first entry)
XX      Probe #6634 used to measure gene expression in human breast sample.
XX      Probe; human; breast disease; breast cancer; development disorder; ss;
XX      inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX      Homo sapiens.
XX      WO200157270-A2.
XX      09-AUG-2001.
XX      29-JAN-2001; 2001WO-US000661.
XX      04-FEB-2000; 2000US-0180312.
XX      26-MAY-2000; 2000US-0207456.
XX      30-JUN-2000; 2000US-0608408.
XX      03-AUG-2000; 2000US-0632366.
XX      21-SEP-2000; 2000US-0234687.
XX      27-SEP-2000; 2000US-0236359.
XX      04-OCT-2000; 2000GB-0024263.
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
XX      WPI; 2001-476286/51.
XX      Novel single exon nucleic acid probe used to measuring gene expression
XX      in a human breast.
XX      Claim 25; SEQ ID No 6634; 322pp; English.
XX      The present invention relates to novel single exon nucleic acid probes.
XX      The present sequence is one such probe. The probes are useful for -
XX      measuring human gene expression in a human breast sample, where the probe
XX      hybridises at high stringency to a nucleic acid expressed in the human
XX      breast. The probes are useful for predicting, diagnosing, grading,
XX      staging, monitoring and prognosing diseases of the human breast,
XX      particularly those diseases with polygenic aetiology. The diseases
XX      include: breast cancer, disorders of development, inflammatory diseases
XX      of the breast, fibrocystic changes, proliferative breast disease and

```

CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other:
 Query Match 26.7%; Score 387; DB 22; Length 567;
 Best Local Similarity 83.7%; Pred. No. 5,2e-117;
 Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
 QY 896 AGGAGTGGGACATACAGTATGACAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 955
 DB 523 AGGACAGTGGGACATACAGTCTTGGCAGTCTCTCTGAGCTCTGCGAGACAGCTCAA 464
 QY 956 CTCACATGATGACCTGGGAGGAGCAAAATGGGAGTTCAATGACATGACACCTGATGA 1015
 DB 463 CTCACAGTGCATACCTGGGAGGAGCAAAAGGAGTTCAATGACATGACACCTGATGA 404
 QY 1016 AGTGGCTCGGCGTTGGGAGAGAGAAAGCAAACTACATGACATGACAACTCAG 1075
 DB 403 GGTGGCCCGGCGTGGGAGAGAGAGAAAGCAAACTACATGACATGACAACTCAG 344
 QY 1076 CGGTGACCTGCTACTACTATGACAAATATATGATGCTAAAGTTGATGATGATG 1135
 DB 343 CGGCGCCCTCGTACTACTATGACAAAGAACATGATGACCAAGGTCATGGAGGCTTA 284
 QY 1136 TGGCTCAAAATTTGATTTCCAGGAAATCGCTCAGAGCCCTCAGAGCTCCAGAAATC 1195
 DB 283 CGGCTCAAAATTTGATTTCCAGGAAATCGCTCAGAGCCCTCAGAGCTCCAGAAATC 224
 QY 1196 ATTCATGATCAAAATTTGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 1255
 DB 223 ATTCATGATCAAAATTTGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 164
 QY 1256 GAGATGAACTTTGATTTGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 1315
 DB 163 GAGATGAACTTTGATTTGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 104
 QY 1316 TGGTGGCCCTCAATGATGATGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 1375
 DB 103 TGGTGGCCCTCAATGATGATGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 44
 QY 1376 GCCAGCTGCTCAATGATGATGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 1418
 DB 43 GCCAGCTGCTCAATGATGATGATTTCCAGGAAATCGCTCAGAGCTCCAGAAATC 1
 RESULT 14
 ABS14215/c
 ID ABS14215 standard; DNA; 567 BP.
 XX
 AC ABS14215;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe ORF from lung SEQ ID No 14206.
 XX
 KM Human; ds; single exon probe; asthma; lung cancer; COPD; IID;
 KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease; open reading frame; ORF.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.

XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 4; SEQ ID No 14206; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other:
 XX
 QY 896 AGGAGTGGGACATACAGTATGACAGTCTCTACTGAGCTTCTGTGGACAGCTCCAA 955
 DB 523 AGGACAGTGGGACATACAGTCTTGGCAGTCTCTCTGAGCTCTGCGAGACAGCTCAA 464

```
QY 956 CTCACATGATCATCCTGGAGAGGCGCAAAATGGGAGTTCAAGATGACAGACCCCGATGA 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 CTCACAGTGTGATTCACCTGGAGAGGAGGAGGAGGAGTTCAAGATGACAGATCCCGAGA 404
QY 1016 AGTGGCTGGGCTGGGAGAGAGAGAAAACCACTAATGAACTATGACAAACCTCAG 1075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 403 GGTGGCCCGGCGCTGGGAGAGAGCGAAGACCAACCCAACTGAATGATGATGAGCTCAG 344
QY 1076 CCGTGCATCTGCTACTACTATGACAAAATATATATGATGATGATGATGATGATGAT 1135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 343 CCGGCGCTCGCTGCTACTACTATGACAAAATATATGATGATGATGATGATGATGATGAT 284
QY 1136 TGCTACAAATTTGATTTCCAGCAATGCTCAGGCTCCAGCCTCCAGCTCCAGATC 1195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 CGCTACAAAGTTGCACTTCCAGGAGATGCGCCAGGCTCCAGCCTCCAGCCTCCAGGAGT 224
QY 1196 ATCCATGACAAATACCATGAGACCTCCCTACATGATGATGATGATGATGATGATGAT 1255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 ATCTCTGACAAAGTACCCCTCAGACCTCCCTACATGATGATGATGATGATGATGATGAT 164
QY 1256 GAAGATGACCTTGTAGCTCCCAATCCCTGCTTTGGCCGTAACCTCATCAGCTTTT 1315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 163 GAAGATGACCTTGTAGCTCCCAATCCCTGCTTTGGCCGTAACCTCATCAGCTTTT 104
QY 1316 TGCTGCCCCATTCATCTGGAATTCACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 103 TGTGCCCCCAAAACCATCTGGAATTCACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 44
QY 1376 GCGAGCTGCTATGCTTCCCTGCTTCCCTGCTTCCCTGCTTCCCTGCTTCCCTGCTT 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 43 CCGCAGCAGCCATATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1

RESULT 15
ABA50193/C
ID ABA50193 standard; DNA: 473 BP.
AC ABA50193;
XX 01-FEB-2002 (first entry)
DE Human breast cell single exon nucleic acid probe #8888.
XX Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX Homo sapiens.
OS WO200157271-A2.
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00662.
PF 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-496933/54.
DR
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes
XX
PS Claim 4; SEQ ID NO 8888; 327pp + sequence listing; English.
```

```
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and Br 474 cells. The method involves contacting
CC the probes with a collection of detectably labeled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
```

Sequence 473 BP; 85 A; 110 C; 167 G; 111 T; 0 other;

Query Match 24.5%; Score 354.6; DB 22; Length 473;

Best Local Similarity 84.4%; Pred. No. 2.4e-106;

Matches 399; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```
QY 897 GGGAGTGGGCGATACACTATGCGAGTCTCTAGTGGAGCTTCTGCGAGAGCTCCAC 956
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 GGGAGTGGGCGATACACTATGCGAGTCTCTAGTGGAGCTTCTGCGAGAGCTCCAC 414
QY 957 TCCAACTGATCACCTGGGAGGCGCAATATGGGAGTTCAAGATGACAGACCTGATGA 1016
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 TCCAGCTGATCACCTGGGAGGCGCAACGAGGAGTTCAAGATGACAGATCCGACGAG 354
QY 1017 GTGGCTGGGCGTTGGGAGAGAGAAAGCAAACTTAATGATGATGATGATGATGATGAT 1076
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 GTGGCTGGGCGTTGGGAGAGAGAAAGCAAACTTAATGATGATGATGATGATGATGAT 294
QY 1077 CGTGACATTCCTACTATGACAAAATATGATGATGATGATGATGATGATGATGATGAT 1136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 CGGCGCTTCCTACTATGACAAAATATGATGATGATGATGATGATGATGATGATGATGAT 234
QY 1137 GCTTACAAATTTGATTTCCAGGAGATGCTAGGCGCTCCAGCTTCCAGCTTCCAGATCA 1196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 GCTTACAAATTTGATTTCCAGGAGATGCTAGGCGCTCCAGCTTCCAGCTTCCAGATCA 174
QY 1197 TCCATGTAACAATACCATGAGACCTCCCTACATGATGATGATGATGATGATGATGAT 1256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 TCTCTGTAACAATACCATGAGACCTCCCTACATGATGATGATGATGATGATGATGAT 114
QY 1257 AAGATGAACTTTGATGCTCCCATTCCTGCTTGGCCGTAACCTCATCAGCTTTT 1316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 AAGATGAACTTTGATGCTCCCATTCCTGCTTGGCCGTAACCTCATCAGCTTTT 54
QY 1317 GCTGCCCCATTCATCTAGTGAATTCACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 53 GCTGCCCCAATTCATCTAGTGAATTCACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1
```

Search completed: July 28, 2003, 11:28:06
Job time : 432.733 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 10:00:14 ; Search time 91.3059 Seconds

(Without alignments)
4860.160 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447
Sequence: 1 gaattccgcgaacgaataat.....gaaagaagcgaagaataa 1447

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCrus.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1447	100.0	1447	US-08-878-177-1	Sequence 1, Appl1
2	1356	93.7	1528	US-08-878-177-3	Sequence 3, Appl1
3	563.8	39.0	2938	US-08-343-443B-3	Sequence 3, Appl1
4	180	12.4	1752	US-09-360-779-1	Sequence 1, Appl1
5	180	12.4	1752	US-09-435-335-1	Sequence 1, Appl1
6	124.6	8.6	1604	US-08-306-691B-43	Sequence 4, Appl1
7	124.6	8.6	1604	PCT-US93-06251-9	Sequence 9, Appl1
8	117.4	8.1	2267	US-09-344-579-1	Sequence 1, Appl1
9	117.4	8.1	2267	US-08-469-412A-1	Sequence 1, Appl1
10	117.4	8.1	2267	US-09-021-715-1	Sequence 1, Appl1
11	115.2	8.0	328	US-08-343-443B-5	Sequence 5, Appl1
12	96	6.6	2064	US-08-875-944B-1	Sequence 1, Appl1
13	96	6.6	2064	US-09-116-049-3	Sequence 3, Appl1
14	95.2	6.6	2410	US-08-780-835B-1	Sequence 1, Appl1
15	95.2	6.6	2410	US-09-303-268-1	Sequence 1, Appl1
16	95.2	6.6	2410	US-09-116-049-1	Sequence 1, Appl1
17	88.6	6.1	2266	US-08-213-767-1	Sequence 1, Appl1
18	87.4	6.0	2544	US-08-469-412A-6	Sequence 6, Appl1
19	87.4	6.0	2544	US-09-021-715-6	Sequence 6, Appl1
20	67.6	4.7	65042	US-09-784-316-3	Sequence 3, Appl1
21	59.8	4.1	852	US-09-020-956-44	Sequence 44, Appl1
22	59.8	4.1	852	US-09-030-607-44	Sequence 44, Appl1
23	59.8	4.1	852	US-09-605-785-44	Sequence 44, Appl1
24	59.8	4.1	852	US-09-439-313-44	Sequence 44, Appl1
25	59.8	4.1	852	US-09-352-616A-44	Sequence 44, Appl1
26	59.8	4.1	852	US-09-232-149A-44	Sequence 44, Appl1
27	58.2	4.0	5427	US-09-009-913-2	Sequence 2, Appl1

28	58.2	4.0	5510	US-09-009-913-3	Sequence 3, Appl1
29	58.2	4.0	5667	US-09-009-913-4	Sequence 4, Appl1
30	55.2	3.8	2975	US-08-368-281-1	Sequence 1, Appl1
31	55.2	3.8	3240	US-08-368-281-3	Sequence 3, Appl1
32	53.4	3.7	1920	US-08-746-789A-1	Sequence 1, Appl1
33	53.4	3.7	1905	US-09-055-113-2	Sequence 2, Appl1
34	48.4	3.3	7218	US-08-232-463-14	Sequence 14, Appl1
35	47.6	3.3	848	US-09-009-913-338	Sequence 338, App
36	47.4	3.3	7218	US-08-232-463-14	Sequence 14, App
37	38.8	2.7	2280	US-09-009-913-8	Sequence 8, Appl1
38	38.8	2.7	2428	US-09-009-913-6	Sequence 6, Appl1
39	38.8	2.7	2498	US-09-009-913-10	Sequence 10, Appl1
40	35.2	2.4	2432	US-08-469-412A-3	Sequence 3, Appl1
41	35.2	2.4	2432	US-09-021-715-3	Sequence 3, Appl1
42	33.4	2.3	1212	US-09-046-578-1	Sequence 1, Appl1
43	32.6	2.3	1447	US-08-878-177-1	Sequence 1, Appl1
44	32.6	2.3	1528	US-08-878-177-3	Sequence 3, Appl1
45	32.6	2.3	12886	US-09-453-702B-14	Sequence 14, Appl1

ALIGNMENTS

```

RESULT 1
US-08-878-177-1
: Sequence 1, Application US/08878177
: Patent No. 6294354
: GENERAL INFORMATION:
: APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: TITLE OF INVENTION: the proteins
: FILE REFERENCE: Chugai Seiyaku Kabushiki Kaisha 5001
: CURRENT APPLICATION NUMBER: US/08/878,177
: CURRENT FILING DATE: 1997-06-18
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO: 1
: LENGTH: 1447
: TYPE: DNA
: ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
US-08-878-177-1

Query Match      100.0%; Score 1447; DB 4; Length 1447;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GAATTCGCGAAGCAATATTTATTTAGCAATTTATTTAGCGATCAATATCTTGATCACA 60
      |||||||
Db      1 GAATTCGCGAAGCAATATTTATTTATTTAGCAATTTATTTAGCGATCAATATCTTGATCACA 60

QY      61 TTATGCGAAGCACTATTATTAAGCAAGCAATTTATTCAGTGTGAGTGAAGACCACTTGTGTTG 120
      |||||||
Db      61 TTATGCGAAGCACTATTATTAAGCAAGCAATTTATTCAGTGTGAGTGAAGACCACTTGTGTTG 120

QY      121 AGTGTCCTGCTGAGTATGCGCCCACTTCCCAAGAAAGAAATGACGCCCTCTCCACAGTG 180
      |||||||
Db      121 AGTGTCCTGCTGAGTATGCGCCCACTTCCCAAGAAAGAAATGACGCCCTCTCCACAGTG 180

QY      181 AATATGCGCAAAATCAATCAAGATGAGCGCGCGCTTCCCAAGAGATGTTATTCACAGC 240
      |||||||
Db      181 AATATGCGCAAAATCAATCAAGATGAGCGCGCGCTTCCCAAGAGATGTTATTCACAGC 240

QY      241 CCCCAGGCAAGTTATCCATTAAGATGAGTGTAAACCAACCAAGTTAATGCGTCAAGA 300
      |||||||
Db      241 CCCCAGGCAAGTTATCCATTAAGATGAGTGTAAACCAACCAAGTTAATGCGTCAAGA 300

QY      301 ATTCACCTGATGACTGAGCGTGGCAAAAGAGGAAATGTTAGAGTTTCAGCAATG 360
      |||||||
Db      301 ATTCACCTGATGACTGAGCGTGGCAAAAGAGGAAATGTTAGAGTTTCAGCAATG 360

QY      361 TTGGAGTGAAGTATGAGAGTACATGAGAAAGCAATATTCGCGCTCCCAATATGACAA 420
      |||||||
Db      361 TTGGAGTGAAGTATGAGAGTACATGAGAAAGCAATATTCGCGCTCCCAATATGACAA 420

```

QY	421	CCAAATGAACGAAAGGTTTTCGTGGCAGAGATNCCAGCTTNTGGAGCCACACACATGAC	480
Db	421	CCAAATGAACGAAAGGTTTTCGTGGCAGAGATNCCAGCTTNTGGAGCCACACACATGAC	480
QY	481	GGCAGCTGGCTGGAGTGGGCACTGAAAGAGATATGGTCTTCCAGAGCTGGACATCTTGTGT	540
Db	481	GGCAGCTGGCTGGAGTGGGCACTGAAAGAGATATGGTCTTCCAGAGCTGGACATCTTGTGT	540
QY	541	TCGAGAACATTTGATGGGAAAGAGTTGTGTAAATATGACCAAGATGACTTCCAGAGACTCA	600
Db	541	TCGAGAACATTTGATGGGAAAGAGTTGTGTAAATATGACCAAGATGACTTCCAGAGACTCA	600
QY	601	CGCCGAGCTATACGACATATATCCCTCGTACACACTACACTACTACTCAGAGAGAGAGAG	660
Db	601	CGCCGAGCTATACGACATATATCCCTCGTACACACTACTACTACTCAGAGAGAGAGAG	660
QY	661	CCACTTTTATTTTTCCAAATACATCAGTTTACCCAGAAAGCAACGCAAGAAATTAACACAA	720
Db	661	CCACTTTTATTTTTCCAAATACATCAGTTTACCCAGAAAGCAACGCAAGAAATTAACACAA	720
QY	721	GGCCGAGATTTACCTTATGAGCAAGAGAGATCAGGCTGGAGCGGTACAGCCATCCCA	780
Db	721	GGCCGAGATTTACCTTATGAGCAAGAGAGATCAGGCTGGAGCGGTACAGCCATCCCA	780
QY	781	CTCAGTCAAAAGCTACCCCAACCATCTTCTTCAAGAGTGCCTCAAAAACAGAAAGCAGCGTC	840
Db	781	CTCAGTCAAAAGCTACCCCAACCATCTTCTTCAAGAGTGCCTCAAAAACAGAAAGCAGCGTC	840
QY	841	CTCAGTTTATTCCTTATCAGATTCCTGGACCGCAGCAGAGCGCGTTCGCAANTCCAGGGA	900
Db	841	CTCAGTTTATTCCTTATCAGATTCCTGGACCGCAGCAGAGCGCGTTCGCAANTCCAGGGA	900
QY	901	GTGGGAGATACAGCTATGGCAGTTTCTTACTGAGAGCTTCTGTGGACAGCTCCAACTCCA	960
Db	901	GTGGGAGATACAGCTATGGCAGTTTCTTACTGAGAGCTTCTGTGGACAGCTCCAACTCCA	960
QY	961	ACTGATACCTGGGAGGGCAACAAATGGGAGTTTCAAGATACAGAGACCTCATAAAGTGG	1020
Db	961	ACTGATACCTGGGAGGGCAACAAATGGGAGTTTCAAGATACAGAGACCTCATAAAGTGG	1020
QY	1021	CTCGGCGCTTGGGAGAGAGAAAAGCAAACTTAACATGAACTATGACAAATCAGCGGTG	1080
Db	1021	CTCGGCGCTTGGGAGAGAGAAAAGCAAACTTAACATGAACTATGACAAATCAGCGGTG	1080
QY	1081	CACTTTCGCTACTACTATGACAAATAATTTATGACTAAAGTTCCATGTTAAACGTATGCGT	1140
Db	1081	CACTTTCGCTACTACTATGACAAATAATTTATGACTAAAGTTCCATGTTAAACGTATGCGT	1140
QY	1141	ACAAATTTGATTTCCAGGAGATGCGCTCAGGCGCTCCAGCGCTCACCCTCCACAAATCATCCA	1200
Db	1141	ACAAATTTGATTTCCAGGAGATGCGCTCAGGCGCTCCAGCGCTCACCCTCCACAAATCATCCA	1200
QY	1201	TGTACAAATATACCCTACAGACTCCCTCAACATGAGATTCCTACATGACACACCCACAGAGA	1260
Db	1201	TGTACAAATATACCCTACAGACTCCCTCAACATGAGATTCCTACATGACACACCCACAGAGA	1260
QY	1261	TGAAGTTTGTATGCTCCCAATCCCGCTGTTTGGCCCGTAACTCATCCAGCTTTTTCCTG	1320
Db	1261	TGAAGTTTGTATGCTCCCAATCCCGCTGTTTGGCCCGTAACTCATCCAGCTTTTTCCTG	1320
QY	1321	CCCCTAATTCATCTGGAATCACCAACTGGAGGCACTACCCCAATACCGAGGCTGCAG	1380
Db	1321	CCCCTAATTCATCTGGAATCACCAACTGGAGGCACTACCCCAATACCGAGGCTGCAG	1380
QY	1381	CTGCTCATATGCTTCCCATCTTGGCACCTTACTAAGTGGGAGAAAGAAAGAAAGCGCC	1440
Db	1381	CTGCTCATATGCTTCCCATCTTGGCACCTTACTAAGTGGGAGAAAGAAAGAAAGCGCC	1440
QY	1441	AAAAAA 1447	
Db	1441	AAAAAA 1447	

```

RESULT 2
US-08-878-177-3
; Sequence 3, Application US/08878177
; Patent No. 6294354
; GENERAL INFORMATION:
; APPLICANT: Chugai Seliyaku Kabushiki Kaisha, 5001, Yamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; TITLE OF INVENTION: the Proteins
; FILE REFERENCE: chugai seliyaku kabushiki kaisha 5001
; CURRENT APPLICATION NUMBER: US/08/878,177
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: C-eryg gene, chicken DNA
US-08-878-177-3

```

Query Match	93.7%	Score 1356	DB 4	Length 1528
Best Local Similarity	94.7%	Pred. No. 0		
Matches 1447	Conservative 0	Mismatches 0	Indels 81	Gaps 1
QY	1	GAATTCGCGCAACGATAATATTATTAATAGCAATTTATAGCATATCAATATCTTGATGACA	60	
DB	1	GAATTCGCGCAACGATAATATTATTAATAGCAATTTATAGCATATCAATATCTTGATGACA	60	
QY	61	TTATGGCAAGACATTTTAAGGAAGCATTTATCAGTGTGTAGTGAACACCAAGTCTTGTTG	120	
DB	61	TTATGGCAAGACATTTTAAGGAAGCATTTATCAGTGTGTAGTGAACACCAAGTCTTGTTG	120	
QY	121	AGTGTGCGCTACGATCGCCGCCACCTGTGCAAAAGACGAATATGACAGCTCTCTTCAGTGTG	180	
DB	121	AGTGTGCGCTACGATCGCCGCCACCTGTGCAAAAGACGAATATGACAGCTCTCTTCAGTGTG	180	
QY	181	AATATGGGCAAAACATCAAAAGATGAGCCGCGCTTCCACAGACATGTGTTATCACAGC	240	
DB	181	AATATGGGCAAAACATCAAAAGATGAGCCGCGCTTCCACAGACATGTGTTATCACAGC	240	
QY	241	CCCCGGCCAGATTTACCTTTAAGATGAGGTGTAAACCCAAACAGTTAATGGGTACAGA	300	
DB	241	CCCCGGCCAGATTTACCTTTAAGATGAGGTGTAAACCCAAACAGTTAATGGGTACAGA	300	
QY	301	ATTTCACCTGATGATCGTCAGCGGTGGCAAAAGAGGGAAATGTTTATGCACTTCAGACAAATG	360	
DB	301	ATTTCACCTGATGATCGTCAGCGGTGGCAAAAGAGGGAAATGTTTATGCACTTCAGACAAATG	360	
QY	361	TTGGGATGACATATGGAAGCTACATGGAAGAGAAACATATTCGCGCTCCAAATATGACAA	420	
DB	361	TTGGGATGACATATGGAAGCTACATGGAAGAGAAACATATTCGCGCTCCAAATATGACAA	420	
QY	421	CCATGGAACGAAGAGTTATTTGTGCGCAGACAGATTCCTACTTTATGGAGCACACACCATGTATC	480	
DB	421	CCATGGAACGAAGAGTTATTTGTGCGCAGACAGATTCCTACTTTATGGAGCACACACCATGTATC	480	
QY	481	GGCAGTGGCTTGAATGGGCACTGGAAGAGTATGGTCTCCAGACGTGACATCTTGTTGT	540	
DB	481	GGCAGTGGCTTGAATGGGCACTGGAAGAGTATGGTCTCCAGACGTGACATCTTGTTGT	540	
QY	541	TCCAGAACATTTGATGGGAAAGAGTTGTGTAAATATGACCAAAAGATGACTTCCAGAGACTCA	600	
DB	541	TCCAGAACATTTGATGGGAAAGAGTTGTGTAAATATGACCAAAAGATGACTTCCAGAGACTCA	600	
QY	601	CGCGGAGCTATTAACGCGACATATCTCTCTGTACACACTTACACTACCTCAGAGAGA-----	654	
DB	601	CGCGGAGCTATTAACGCGACATATCTCTCTGTACACACTTACACTACCTCAGAGAGACTCTCTC	660	
QY	655	-----GAGGAGCCCACTTTTATTTTTCCAAAATACATCAAGTTATACCGAGAG	699	
DB	661	TTCCACATTTGACTTTCAGATGATGTGTATTAAGGCGCTTACAAACCTCTCCACGGTTAATGC	720	
QY	655	-----GAGGAGCCCACTTTTATTTTTCCAAAATACATCAAGTTATACCGAGAG	699	


```

DB          426 TTCACTGACAGTTCATTTTCACAACTGGAGTGGTCAATTTACC 472

RESULT 10
US-09-021-715-1
; Sequence 1, Application US/09021715
; Patent No. 6194547
GENERAL INFORMATION:
APPLICANT: MAVROTHALASSITIS, George J.
          Bialf, Donald G.
          Fisher, Robert J.
          Beal Jr., Gregory J.
          Athanasious, Merope A.
          Souras, Dionysios N.
TITLE OF INVENTION: The ERF genetic Locus and Its Products
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,715
FILING DATE: 10-Feb-1998
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 015280-229000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2667 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 123..1769
OTHER INFORMATION: /note= "human ERF (ETS2 Repressor
Factor) cDNA"
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-021-715-1

Query Match      8.1%; Score 117.4; DB 4; Length 2667;
Best Local Similarity 63.1%; Pred. No. 8.8e-30;
Matches 181; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

DB          1074 AGCGGTGACACTTCCTTACTATGACAAAATATATGACTAAAGTTGATGTAACGC 1133
DB          366 AGCGGTGACACTTCCTTACTATGACAAAATATATGACTAAAGTTGATGTAACGC 425

```

0Y 1134 ATGSCCAGCAATTTGATTTCACGGAAATCGCTCAGGCCCTCCAGGC 1180
| | | | | | | | | | | | | | | | | | | | | |
DB 426 TTACACTACAGTTCATTCAATTAACAACAACTGTGCTGTCATTACC 472

RESULT 11
US-08-343-443B-5
; Sequence 5, Application US/08343443B
; Patent No. 598734
GENERAL INFORMATION:
APPLICANT: Aurias, Alain
APPLICANT: Delattre, Olivier
APPLICANT: Desmazes, Chantal
APPLICANT: Melot, Thomas
APPLICANT: Peter, Martine
APPLICANT: Ploougastel, Beatrice
APPLICANT: Thomas, Gilles
APPLICANT: Zucman, Jessica
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEOTIC ACIDS OF FUSION RESULTING FROM SAID
TRANSLATIONS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCY/FR93/00494
FILING DATE: 19-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989,6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 328 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..327
US-08-343-443B-5

Query Match 8.0%; Score 115.2; DB 2; Length 328;
Best Local Similarity 67.4%; Pred. No. 1.2e-29;
Matches 178; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

0Y 733 CTATATGACGACGACGAGCATCAGCGCTGGACGACGATCACGCCATCC---CACTCAGTCAA 789
| | | | | | | | | | | | | | | | | | | | | |
DB 65 CTTATGATCTCAGTCAGAACAGGAGCTTGGGGCAATAAATGAATTCCTGGCCCTCAACAAA 124

OY 790 AAGTACCCACATCTTCAACAGTGGCCCAAGAGACGAGCTCTCATGTTAG 849
Db 125 GTCTCTCCCTTGAGGGGACACAGTACAGTAAGTAATACAGCAACGAGCCCGAGCAG 184
OY 850 ATCTTATACAGATCTTGGACCGACGACGAGCTTGTGAATTCACGGAGTGGGAGA 909
Db 185 ATCCGTATACAGATCTTGGACCGACGACGAGCTTGTGAATTCACGGAGTGGGAGA 244
OY 910 TACAGCTATGACATCTTCTACTGAGCTTGTGTGAGACAGCTCCCACTCAATGATCA 969
Db 245 TCCAGCTGTGACATCTTCTCTGTGAGCTGTCTCCGACAGCGCCAGCGAGCTATCA 304
OY 970 CCTGGAGGCGACAAATGGGAGT 993
Db 305 CCTGGAGGCGACCAACGGGAGT 328

RESULT 12

US-08-875-944B-1
Sequence 1, Application US/08875944B
Patent No. 6096542
GENERAL INFORMATION:
APPLICANT: FUJINAGA, Kei
APPLICANT: YOSHIDA, Koichi
APPLICANT: HIGASHINO, Fumihito
TITLE OF INVENTION: CANCER CONTROL
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 624 Ninth Street N.W., Ste. 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,944B
FILING DATE: 07-AUG-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 07-020173
FILING DATE: 08-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/00016
FILING DATE: 09-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: FUJINAGA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2064 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1386
US-08-875-944B-1

Query Match

Best Local Similarity 62.5%; Score 96; DB 3; Length 2064;
Matches 150; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

OY 910 TACAGCTATGACATCTTCTACTGAGCTTGTGTGAGACAGCTCCCACTCAATGATCA 969

Db 956 TGCAGCTGTGCAATTTCTGTGCTGCTTGATGACCCAAACAAATGCCATTTCAATG 1015
OY 970 CTTGGAGAGGCAAAATGGGAGTTTCAAGATGACAGACCCGATGAAATGGCTGGCGGT 1029
Db 1016 CTTGGAGAGGCGCGGGAATGAGTTCAAGCTTCAATGACCTTGAGAGAGTCCGAGGCTCT 1075
OY 1030 GGGGAGAGAGGAAAGCAAACTTAACATGAACTATGACAACTGACGCGTGCCTTGCCT 1089
Db 1076 GGGGATTCACAGAAACCGCGCATGATTAAGAACTGAGAGCGCTGCTCCGAT 1135
OY 1090 ACTACTATGACAAAATTTATGACTTAAGTTGATGTAACGCTATGCTTACAAATTTG 1149
Db 1136 ACTATTTATGAAAGGACATATGCAAGAGTGGCTGTGACGCTTACGTATCAAGTTTG 1195

RESULT 13

US-09-116-049-3
Sequence 3, Application US/09116049A
Patent No. 6248351
GENERAL INFORMATION:
APPLICANT: Hung, Men-Chie
TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
FILE REFERENCE: UTSC:582
CURRENT APPLICATION NUMBER: US/09/116,049A
CURRENT FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 3
LENGTH: 2064
TYPE: DNA
ORGANISM: Homo sapiens
US-09-116-049-3

Query Match

Best Local Similarity 62.5%; Score 96; DB 4; Length 2064;
Matches 150; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

OY 910 TACAGCTATGACATCTTCTACTGAGCTTGTGTGAGACAGCTCCCACTCAATGATCA 969
Db 956 TGCAGCTGTGCAATTTCTGTGCTGCTTGATGACCCAAACAAATGCCATTTCAATG 1015
OY 970 CTTGGAGAGGCAAAATGGGAGTTTCAAGATGACAGACCTGATGAAGTGGCTGGCGGT 1029
Db 1016 CTTGGAGAGGCGCGGGAATGAGTTCAAGCTTCAATGACCTTGAGAGAGTGGCGGCTCT 1075
OY 1030 GGGGAGAGAGGAAAGCAAACTTAACATGAACTATGACAACTGACGCGTGCCTTGCCT 1089
Db 1076 GGGGATTCACAGAAACCGCGCATGATTAAGAACTGAGAGCGCTGCTCCGAT 1135
OY 1090 ACTACTATGACAAAATTTATGACTTAAGTTGATGTAACGCTATGCTTACAAATTTG 1149
Db 1136 ACTATTTATGAAAGGACATATGCAAGAGTGGCTGTGAGCGTTAGCTTACAGATTTG 1195

RESULT 14

US-08-780-835B-1
Sequence 1, Application US/08780835B
Patent No. 5922688
GENERAL INFORMATION:
APPLICANT: Hung, Men-Chie
APPLICANT: Xing, Xiangming
TITLE OF INVENTION: PEA3 is a Tumor Suppressor
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,835B
FILING DATE: 10-JAN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: UTSC500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2410 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-780-835B-1

Query Match 6.6%; Score 95.2; DB 2; Length 2410;
Best Local Similarity 57.9%; Pred. No. 3.9e-22;
Matches 169; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 865 TTGACCGACGACGACCGCTTTGCAATCCAGGAGTGGGCAATACAGCTATGGCAGT 924
DB 1315 TTGAGCTTTCCGGAGGGGACCCCTACACGCGCGGGGTGCTTACAACTGTGGCAGT 1374
QY 925 TCTACTGAGCTTCTGCGACAGCTCCAACTGCATCATCCTCGGAGGAGGACAA 984
DB 1375 TTGAGGTGCGCCCTCTGATATACCCCAAAATGCTCATTTTCATTGCTTGGACAGCGCGG 1434
QY 985 ATGGGAGTTCAAGATGACAGACCCCTGATGAGTGGCTCGGCTTGGGAGAGAGAAA 1044
DB 1435 GAATGAGTTAACTAATTAATGACAGCTGAGAGTGGCGCTCTGCGATTCAGAGAA 1494
QY 1045 GCAACCTAACATGACATGACAACTCAGCCGCTGCACCTCGCTACTACTATGACAAA 1104
DB 1495 ACCGCGCCAGCATTAATATGACAGCTGAGCGCGCTCGCTGCACTACTATTATGAGAA 1554
QY 1105 ATATTATGACTAAGTTCTATGTAACGCTATGCTTACAAATTTGATTCCA 1156
DB 1555 GCATCATGACAGAGGTGGCTGGCGAAGCCTACGTTGATCAAGTTGTGTGCGA 1606

RESULT 15
US-09-303-268-1

; Sequence 1, Application US/09303268
; Patent No. 6172212

; GENERAL INFORMATION:

; APPLICANT: Hung, Mien-Chie

; Xing, Xiangming

; TITLE OF INVENTION: PEAS is a Tumor Suppressor

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARNOLD, WHITE AND DURKEE

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210-4433

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/303,268

; FILING DATE: 30-Apr-1999

; CLASSIFICATION: <Unknown>

APPLICATION NUMBER: 08/780,835
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: UTSC500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2410 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-303-268-1

Query Match 6.6%; Score 95.2; DB 4; Length 2410;
Best Local Similarity 57.9%; Pred. No. 3.9e-22;
Matches 169; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 865 TTGACCGACGACGACCGCTTTGCAATCCAGGAGTGGGCAATACAGCTATGGCAGT 924
DB 1315 TTGAGCTTTCCGGAGGGGACCCCTACACGCGCGGGGTGCTTACAACTGTGGCAGT 1374
QY 925 TCTACTGAGCTTCTGCGACAGCTCCAACTGCATCATCCTCGGAGGAGGACAA 984
DB 1375 TTGAGGTGCGCCCTCTGATATACCCCAAAATGCTCATTTTCATTGCTTGGACAGCGCGG 1434
QY 985 ATGGGAGTTCAAGATGACAGACCCCTGATGAGTGGCTCGGCTTGGGAGAGAGAAA 1044
DB 1435 GAATGAGTTAACTAATTAATGACAGCTGAGAGTGGCGCTCTGCGATTCAGAGAA 1494
QY 1045 GCAACCTAACATGACATGACAACTCAGCCGCTGCACCTCGCTACTACTATGACAAA 1104
DB 1495 ACCGCGCCAGCATTAATATGACAGCTGAGCGCGCTCGCTGCACTACTATTATGAGAA 1554
QY 1105 ATATTATGACTAAGTTCTATGTAACGCTATGCTTACAAATTTGATTCCA 1156
DB 1555 GCATCATGACAGAGGTGGCTGGCGAAGCCTACGTTGATCAAGTTGTGTGCGA 1606

Search completed: July 28, 2003, 12:16:34
Job time : 93.3059 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	length	DB	ID	Description
1	144.7	100.0	144.7	11	US-09-902-772-1	Sequence 1, April
2	135.6	93.7	152.8	11	US-09-902-772-3	Sequence 3, April
3	92.1	63.6	31.6	15	US-10-205-823-98	Sequence 98, April
4	38.7	26.7	15.7	10	US-09-864-761-18410	Sequence 18410, April
5	354.6	24.5	4.73	10	US-09-864-761-20472	Sequence 20472, April
6	30.7	21.2	4.67	12	US-09-918-995-23356	Sequence 23356, April
7	276.6	19.1	5.49	10	US-09-922-779-54	Sequence 54, April
8	25.5	17.6	4.20	10	US-09-864-761-5705	Sequence 3705, April
9	25.5	17.6	4.54	10	US-09-864-761-1653	Sequence 1653, April
10	180	12.4	1.52	10	US-09-850-799-1	Sequence 1, April
11	167.8	11.6	5.53	10	US-09-864-761-7427	Sequence 7427, April
12	163.2	11.3	3.48	10	US-09-864-761-29364	Sequence 29364, April
13	163.2	11.3	4.77	10	US-09-864-761-5035	Sequence 5035, April
14	163.2	11.3	4.78	10	US-09-864-761-12798	Sequence 12798, April
15	162.2	11.2	2.25	10	US-09-864-761-21761	Sequence 21761, April
16	148.4	10.3	2.33	15	US-10-037-270-869	Sequence 869, April

17	122.8	8.5	416	12	US-09-918-995-8472	Sequence 8372, A
18	122.4	8.5	165	10	US-09-864-761-33294	Sequence 23294, A
19	122.2	8.4	473	12	US-09-918-995-27904	Sequence 27904, A
20	122.2	8.4	1884	11	US-09-925-300-420	Sequence 420, App
21	122.2	8.4	2188	11	US-09-920-300A-1716	Sequence 1716, App
22	122.2	8.4	2188	15	US-10-033-528-1716	Sequence 1716, App
23	122.2	8.4	2268	11	US-09-920-300A-1693	Sequence 1693, App
24	122.2	8.4	2268	15	US-10-033-528-1693	Sequence 1693, App
25	122.2	8.4	2269	11	US-09-954-551-955	Sequence 955, App
26	122.2	8.4	3652	15	US-10-106-658-824	Sequence 824, App
27	115.8	8.0	3178	15	US-10-108-605-124	Sequence 124, App
28	115.8	8.0	3178	15	US-10-108-605-128	Sequence 128, App
29	113	7.8	2952	10	US-09-925-297-257	Sequence 257, App
30	113	7.8	406	15	US-10-084-817-187	Sequence 187, App
31	101	7.0	300	15	US-10-177-063-15	Sequence 15, App
32	99.4	6.9	300	15	US-10-177-063-16	Sequence 16, App
33	96	6.6	2064	12	US-09-884-363-3	Sequence 3, App
34	96	6.6	2064	15	US-10-167-485-1	Sequence 1, App
35	96	6.6	2333	11	US-09-920-300A-1788	Sequence 1788, App
36	96	6.6	2333	11	US-09-880-107-5316	Sequence 3316, App
37	96	6.6	2333	15	US-10-033-528-1788	Sequence 1788, App
38	95.2	6.6	2410	12	US-09-884-363-1	Sequence 1, App
39	90.6	6.3	2256	15	US-10-036-841-17	Sequence 317, App
40	75.6	5.2	180	15	US-10-177-063-27	Sequence 27, App
41	70.4	4.9	4190	11	US-09-962-832-111	Sequence 111, App
42	67.6	4.7	65042	15	US-10-229-124-3	Sequence 3, App
43	66.2	4.6	414	10	US-09-864-761-10426	Sequence 10426, A
44	65.8	4.5	303	10	US-09-864-761-2705	Sequence 2705, A
45	63.2	4.4	496	12	US-09-918-995-6907	Sequence 6907, App

ALIGNMENTS

RESULT 1
 US-09-902-772-1
 : Sequence 1' Application US/09902772
 : Patent No. US20020164739a1
 : GENERAL INFORMATION:
 : APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
 : TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
 : TITLE OF INVENTION: the Proteins
 : FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
 : CURRENT APPLICATION NUMBER: US/09/902,772
 : PRIOR FILING DATE: 2001-07-12
 : PRIOR APPLICATION NUMBER: US/08/878,177
 : PRIOR FILING DATE: 1997-06-18
 : NUMBER OF SEQ. ID NOS: 7
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 1447
 : TYPE: DNA
 : ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
 : US-09-902-772-1

	Query Match	Similarity	100.0%	Score 1447	DB 11	Length 1447
	Best Local	Similarity	100.0%	Pred. No. 0		
	Matches 1447	Conservative	0	Mismatches	0	Indels
					0	Gaps
					0	
QY	1	GAATTCGGGAAAGAAATTAATTTATTTATGCAATTTATGCGATCAATTAATCTTGATCACA	60			
Db	1	GAATTCGGGAAAGAAATTAATTTATTTATGCAATTTATGCGATCAATTAATTAATCTTGATCACA	60			
QY	61	TTATGCGACAGCACTTTAAGCAAGCATTTATGAGTGAGTGAAGACAGCATCTCTGTTTG	120			
Db	61	TTATGCGACAGCACTTTAAGCAAGCATTTATGAGTGAGTGAAGACAGCATCTCTGTTTG	120			
QY	121	AGGTGCTTCAGCGGATTCGCCCACTTGTGGAAGACAGAAATATACAGCTCTCTTCCAGTG	180			
Db	121	AGGTGCTTCAGCGGATTCGCCCACTTGTGGAAGACAGAAATATACAGCTCTCTTCCAGTG	180			
QY	181	AATATGGGAAACATCAAAAGATGAGCCGCCGCGTTCGCCAGCAGACTGTGTTATCAGC	240			
Db	181	AATATGGGAAACATCAAAAGATGAGCCGCCGCGTTCGCCAGCAGACTGTGTTATCAGC	240			

```

Db 181 AATATGGGCAAAACATCAAGATGAGCCGGCGCTTCCCAAGACGAGACTGGTTATCACAGC 240
QY 241 CCCCCGCGACAGTATACCATTAAGATGAGTGTATACCAACCAAGTTAATGGGTCAGGA 300
Db 241 CCCCCGCGACAGTATACCATTAAGATGAGTGTATACCAACCAAGTTAATGGGTCAGGA 300
QY 301 ATTACCTGATGACTGCGAGGTGGCAAAAGAGGAAATGGTTAGACAGTTCAACATATG 360
Db 301 ATTACCTGATGACTGCGAGGTGGCAAAAGAGGAAATGGTTAGACAGTTCAACATATG 360
QY 361 TTGGGATGAGACTATGAGAGCTACATGAGAGAAACATATTCGGCTCCAAATATGACAA 420
Db 361 TTGGGATGAGACTATGAGAGCTACATGAGAGAAACATATTCGGCTCCAAATATGACAA 420
QY 421 CCATGTAACGAGAGTATTTGGCCAGACATCTAGTTATGGAGACAGACCATGTATAC 480
Db 421 CCATGTAACGAGAGTATTTGGCCAGACATCTAGTTATGGAGACAGACCATGTATAC 480
QY 481 GGCAGTGGCTGAGTGGGAGTGAAGAGATATGGTCTTCCAGACGTGGACATCTTGTGT 540
Db 481 GGCAGTGGCTGAGTGGGAGTGAAGAGATATGGTCTTCCAGACGTGGACATCTTGTGT 540
QY 541 TCCGAAACATTTGATGGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCCAGACTCA 600
Db 541 TCCGAAACATTTGATGGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCCAGACTCA 600
QY 601 CGCGGAGCTATTAACGAGATATCTCTCTGTCACACTACACTACTCTAGAGAGAGAGAG 660
Db 601 CGCGGAGCTATTAACGAGATATCTCTCTGTCACACTACACTACTCTAGAGAGAGAGAG 660
QY 661 CCACCTTTATTTTCCAAATACATACATTTACCAGAGAGACGCAAGCAAGATTAACAACAA 720
Db 661 CCACCTTTATTTTCCAAATACATACATTTACCAGAGAGACGCAAGCAAGATTAACAACAA 720
QY 721 GGCAGATTTACCTTATGAGACAGGAGAGATAGCGTGGAGACATGACACCATTCCTCA 780
Db 721 GGCAGATTTACCTTATGAGACAGGAGAGATAGCGTGGAGACATGACACCATTCCTCA 780
QY 781 CTCAGTCAAAAAGCTACCCCAACCATCTTCAACAGTGGCCCAAAAGAGAGAGAGAGGTC 840
Db 781 CTCAGTCAAAAAGCTACCCCAACCATCTTCAACAGTGGCCCAAAAGAGAGAGAGAGGTC 840
QY 841 CTCAGTGAATCTTATCAAGATCTTGTGAGCGACAGACAGCGCTTTCGCAATCCAGAGGA 900
Db 841 CTCAGTGAATCTTATCAAGATCTTGTGAGCGACAGACAGCGCTTTCGCAATCCAGAGGA 900
QY 901 GTGGGCGAGATACACATTTGAGAGTCTTCTAGTGAGCTTGTGTGGAGAGTCCCAATCCCA 960
Db 901 GTGGGCGAGATACACATTTGAGAGTCTTCTAGTGAGCTTGTGTGGAGAGTCCCAATCCCA 960
QY 961 ACTGCATCACCTGGAGAGGAGCAAAATGGGAGTTCAGATGACAGACCTGTATGATGG 1020
Db 961 ACTGCATCACCTGGAGAGGAGCAAAATGGGAGTTCAGATGACAGACCTGTATGATGG 1020
QY 1021 CTGGGCGTGGGAGAGAGAGAAAGCAACCTTAACATGAATGACAAACTCAGCGGTG 1080
Db 1021 CTGGGCGTGGGAGAGAGAGAAAGCAACCTTAACATGAATGACAAACTCAGCGGTG 1080
QY 1081 CACTTGGCTACTACTATGACAAAATTTATGACTTAAGTTCAATGATGAAGCGTATCCCT 1140
Db 1081 CACTTGGCTACTACTATGACAAAATTTATGACTTAAGTTCAATGATGAAGCGTATCCCT 1140
QY 1141 ACAAAATTTGATTTCCAGAGATGCTCAGAGCCCTCAGGCTCAACCTCCAGAAATCATCCA 1200
Db 1141 ACAAAATTTGATTTCCAGAGATGCTCAGAGCCCTCAGGCTCAACCTCCAGAAATCATCCA 1200
QY 1201 TGTACAAATATCCATCAGACCTCCCTCAACATGAGTCTTCAATGACACACCCCAAGAGA 1260
Db 1201 TGTACAAATATCCATCAGACCTCCCTCAACATGAGTCTTCAATGACACACCCCAAGAGA 1260
QY 1261 TGAACCTTTGAGTCCCATCCCGCTCTTGGCCGGAATCATCAGCTTTTGTCTG 1320
Db 1261 TGAACCTTTGAGTCCCATCCCGCTCTTGGCCGGAATCATCAGCTTTTGTCTG 1320

```

```

QY 1321 CCCCTAATCCATACTGGAATTCACCAACTGGAGGACATCTACCCCAATACGAGGCTGCAG 1380
Db 1321 CCCCTAATCCATACTGGAATTCACCAACTGGAGGACATCTACCCCAATACGAGGCTGCAG 1380
QY 1381 CTGCTCATATGACCTTCCCATCTTGGACACTACTACTAAGTGGGAGAAAGAAAGAGCGCC 1440
Db 1381 CTGCTCATATGACCTTCCCATCTTGGACACTACTACTAAGTGGGAGAAAGAAAGAGCGCC 1440
QY 1441 AAGAAAA 1447
Db 1441 AAGAAAA 1447

RESULT 2
US-09-902-772-3
; Sequence 3, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification suppressing Proteins and Genes of
; TITLE OF INVENTION: the Proteins
; FILE REFERENCE: Chugai Seiyaku Kabushiki Kaisha 5001
; CURRENT APPLICATION NUMBER: US/09/902.772
; PRIOR FILING DATE: 2001-07-12
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: c-ery gene, chicken DNA
US-09-902-772-3

Query Match 93.7%; Score 1356; DB 11; Length 1528;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1447; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 1 GAATTCGGGCAAGCAATTAATTTATGCAATTTATAGCAATTTATAGGCAATTAATCTGATCACA 60
Db 1 GAATTCGGGCAAGCAATTAATTTATGCAATTTATAGCAATTTATAGGCAATTAATCTGATCACA 60
QY 61 TTATGCGAACACATTAATTAAGGAAGCATTTACAGTGTGATGAGAGACAGTCCCTGTTTG 120
Db 61 TTATGCGAACACATTAATTAAGGAAGCATTTACAGTGTGATGAGAGACAGTCCCTGTTTG 120
QY 121 AGTGTGCTAAGGATCGCCCGACCTTGCAAAAGACAGAAATGACAGCTCTCTTCCAGTGG 180
Db 121 AGTGTGCTAAGGATCGCCCGACCTTGCAAAAGACAGAAATGACAGCTCTCTTCCAGTGG 180
QY 181 AATATGGGCAAAACATCAAAAGATGAGCGCGGCGTTCGCCAGAGAGACTGTTATACAGC 240
Db 181 AATATGGGCAAAACATCAAAAGATGAGCGCGGCGTTCGCCAGAGAGACTGTTATACAGC 240
QY 241 CCCCCGCGAGAGTACCATTAAGATGAGTGAAGAGAGGAGGAAATGGTTAGGTTCAAGGA 300
Db 241 CCCCCGCGAGAGTACCATTAAGATGAGTGAAGAGAGGAGGAAATGGTTAGGTTCAAGGA 300
QY 301 ATTACCTGATGACTGCGAGGTGGCAAAAGAGGAGGAAATGGTTAGGTTCAAGCATG 360
Db 301 ATTACCTGATGACTGCGAGGTGGCAAAAGAGGAGGAAATGGTTAGGTTCAAGCATG 360
QY 361 TTGGGATGAGACTATGAGAGCTACATGAGAGAAAGCAATTCGGCGCTCCAAATATGACAA 420
Db 361 TTGGGATGAGACTATGAGAGCTACATGAGAGAAAGCAATTCGGCGCTCCAAATATGACAA 420
QY 421 CCAATGAGCAAGAGTATTTGTGCCAGACAGATCTTACGTTATGAGAGACAGACATGTAC 480
Db 421 CCAATGAGCAAGAGTATTTGTGCCAGACAGATCTTACGTTATGAGAGACAGACATGTAC 480
QY 481 GGCAGTGGCTGAGAGTGGGAGTGAAGAGATGATGCTTCCAGACGTGACATCTTGTGT 540
Db 481 GGCAGTGGCTGAGAGTGGGAGTGAAGAGATGATGCTTCCAGACGTGACATCTTGTGT 540

```



```

OY 554 TGGGAAGAGTTGTGTAATAATGACCAAGATGATCTCCAGAGACTCAGCCGAGTATAA 613
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 769 TGGGAAGAGTGTGTAAGATGACCAAGAGACTTCCAGAGGCTCAGCCAGTACAA 828
OY 614 CGCAGATATCTCTGTGCACACCTACCTACCTCAGAGAGA -----GAGGAGCCAC 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 829 CGCCGATCTCTCTCAGATCTCAGATCTCAGAGAGACTCTCTTCCACATTTGAC 888
OY 665 TTTATTTTCCAAATACATACGTTTACCCGAGCAAGCAAGAAAGTAAACAAAGGCC 724
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 889 TTCAGATGATGTTGATTAAGGCTTACAAACTCTCCAGGTTAATGATGCTTAACAAAC 948
OY 725 AGATTATCTTATGAGCAAGGAGATGAGCTGAGCTGAGCTGAGCTTACCACTTCCCA 781
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 949 AGATTATCTTATGAGCAAGGAGATGAGCTGAGCTGAGCTGAGCTTACCACTTCCCA 1008
OY 782 TCAGTCAAAAGCTACCAACCATCATCTTCAACAGTGCCCAAAACAGAGACGCTCC 841
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1009 CCACTGCAAAAGCTCTCAACCATCATCTTCCACAGTGCCCAAAACTGAAAGACGCTCC 1068
OY 842 TCAGTATGATCTTATGATGATCTTGGACGACGACGAGCTTGTGAATCCAGGAG 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1069 TCAGTATGATCTTATGATGATCTTGGACGACGACGAGCTTGTGAATCCAGGAG 1128
OY 902 TGGGAGATACAGCTATGAGCTTCTACTGAGCTTCTGTGGAGAGCTTCACTCCAA 961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1129 TGGGAGATACAGCTTGTGGAGCTTCTCTGAGCTTCTGTGGAGAGCTTCACTCCAG 1188
OY 962 CTGATCATCCTGAGGAGGCAAAATGGGAGCTTCAAGTACAGACCTGTATGAAGTGC 1021
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1189 CTGATCATCCTGAGGAGGCAAAAGGAGGAGTCAAGTACAGGAGCTTCAAGTGC 1248
OY 1022 TGGGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1249 CGGAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1308
OY 1082 ACTTGGTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1309 CCTCGTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1368
OY 1142 CAATTTGATTTCCACGGAATCGCTCAGGCTTCCAGCTTCCAGCTTCCAGATATCAT 1201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1369 CAATTTGATTTCCACGGAATCGCTCAGGCTTCCAGCTTCCAGCTTCCAGATATCAT 1428
OY 1202 GTACAAATACCATCAGACCTCCCTCATAGTTCCTACATCAGACCCCGAGAGAT 1261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1429 GTACAAATACCATCAGACCTCCCTCATAGTTCCTACATCAGACCCCGAGAGAT 1488
OY 1262 GAACCTTGTAGCTCCCATCCCTGCTTCCCTGTAACCTCATCAGCTTTTGTGTC 1321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1489 GAACCTTGTAGCTCCCATCCCTGCTTCCCTGTAACCTCATCAGCTTTTGTGTC 1548
OY 1322 CCTCAATCATCTGGAATTCACCACTGAGGAGCTTACCCCAATACCAAGCTGCGAC 1381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1549 CCTCAATCATCTGGAATTCACCACTGAGGAGCTTACCCCAATACCAAGCTGCGAC 1608
OY 1382 TGCTCATATGCTTCCCATCTTGGACACTTACTATA 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1609 CAGCATATGCTTCTCATCTGGGACTTACTATA 1645

```

```

RESULT 4
US-09-864-761-18410/c
: Sequence 18410, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

```

: FILE REFERENCE: Aecomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 18410
: LENGTH: 567
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AP000163.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.92
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
: OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 9.00e-99
: OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00
: OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00
: US-09-864-761-18410

```

```

Query Match 26.7% Score 387 DB 10 Length 567
Best Local Similarity 83.7% Pred. No 8.5e-117
Matches 438 Conservative 0 Mismatches 85 Indels 0 Gaps 0
OY 896 AGGAGTGGGAGATGAGCTATGCGAGTTCCTACTGAGCTTGTGGAGAGTCCAA 955
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 523 AGGAGTGGGAGATGAGCTATGCGAGTTCCTACTGAGCTTGTGGAGAGTCCAA 464
OY 956 CTCCACTGATCATCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 CTCCACTGATCATCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 404

```

```

RESULT 5
US-09-864-761-20472/c
Sequence 20472, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Gene Expression Analysis by Microarray
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670

```

Query Match	24.5%	Pred. 354.6;	DB 10;	Length 473;
Best Local Similarity	84.4%;	Pred. No. 3.8e-106;		
Matches 399;	Conservative 0;	Mismatches 74;	Indels 0;	Gaps 0;

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FROM VARIOUS CDNA LIBRARIES
 ; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3705
LENGTH: 420
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000021.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
US-09-864-761-3705

Query Match 17.6%; Score 255; DB 10; Length 420;
Best Local Similarity 81.9%; Pred. No. 2.8e-73;

Matches 294; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1060 ACTATGCAAACTACGCGCTGCACTTGTCTACTACTATGCAAAATATTATGCTAAAG 1119
DB 420 ACTACGATTAAGCTCAGCGCGGCGCTCGTACTACTACTATGCAAGAAATCATGACCAAG 361
QY 1120 TTATGGTAAACGCTATGCTTACAAATTTGATTTCCAGGAATGCTGAGCCCTCCAGC 1179
DB 360 TCATGGGAAGCGGTAGCGCTACAAATTCGACTCCAGGAGTCCGAGGCGCTCCAGC 301
QY 1180 CTCACCTCCAGATCATTCATGCAATACCATGAGACTCCCTACATGAGTTCCT 1239
DB 300 CCCACCCCGGAGCTCTCTCTGACAAAGTACCCCTGAGACTCCCTACATGAGGCTCCT 241
QY 1240 ACCATGCAACCCCGGAGATGATGAACTTTGTAGTCCCATCCCTGCTTTGCGCGTAA 1299
DB 240 ATACGCGCAACCCGAGAAATGAACTTTGTGGGGCCCAACCTCCAGCCCTCCGCTGA 181
QY 1300 CCTCATCAGACTTTTGTGCTGCGCCCTATCATCTAGTGAATTCACCACTGAGGCACT 1359
DB 180 CACTCTCCACTTTTGTGCTGCGCCCAACCACTAGTGAATTCACCACTGAGGCACTAT 121
QY 1360 ACCCGAATACGAGGCTGCGCTGCTCATATGCTTCCCATTTGGAGCTACTACTAA 1418
DB 120 ACCCGAATACGAGGCTGCGCTGCGCTCATATGCTTCCCATTTGGAGCTACTACTAA 62

RESULT 9

US-09-864-761-1653/c
Sequence 1653, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Neomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 1653
LENGTH: 454
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000163.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
US-09-864-761-1653

Query Match 17.6%; Score 255; DB 10; Length 454;
Best Local Similarity 81.9%; Pred. No. 2.9e-73;

Matches 294; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 1060 ACTATGCAAACTACGCGCTGCACTTGTCTACTACTATGCAAAATATTATGCTAAAG 1119
DB 454 ACTACGATTAAGCTCAGCGCGGCGCTCGTACTACTACTATGCAAGAAATCATGACCAAG 395
QY 1120 TTATGGTAAACGCTATGCTTACAAATTTGATTTCCAGGAATGCTGAGCCCTCCAGC 1179
DB 394 TCATGGGAAGCGGTAGCGCTACAAATTCGACTTCCAGGAGTTCGCCAGCGCTCCAGC 335
QY 1180 CTCACCTCCAGATCATTCATGCAATACCATGAGACTCCCTACATGAGTTCCT 1239
DB 334 CCCACCCCGGAGTCTCTCTGACAAAGTACCCCTGAGACTCCCTACATGAGGCTCCT 275

QY	1240	ACCATGACACACCCAGAAAGTGAACCTTTGAGTGTCCCAATCCCGCGCTTGGCCGAA	1299
Db	274	ATACGCGCCACCCACAGAAAGTGAACCTTTGAGTGTCCCAATCCCGCGCTTGGCCGAA	215
QY	1300	CCATCATCAGACTTTTGTGCGCCCTAATGCATCTGGAATATCAACATGGAGGACATCT	1359
Db	214	CATCTTCAGTTTTTTTGTGCTGTGCCCCAATACCATATCTGGAATATCAACATGGAGGATAT	155
QY	1360	ACCCCAATACAGGCTGCCAGAGCTCATATATGCTTCCCATCTTGGGACACCTAATCTAA	1418
Db	154	ACCCCAACACTAGGCTCCCAACACAGCATATGCTTTCATCTGTGGGACTTACTACTAA	96

```

RESULT 10
US-09-850-799-1
Sequence 1, Application US/09850799
Patent No. US20020090647A1
GENERAL INFORMATION:
APPLICANT: Deneris, Evan S.
APPLICANT: Eyodoro, Dmitry V.
APPLICANT: Hendricks, Timothy J.
TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
FILE REFERENCE: Case-03828
CURRENT APPLICATION NUMBER: US/09/850,799
CURRENT FILING DATE: 2001-05-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/360,779
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-26
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1752
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: CDS
LOCATION: (112)..(1131)
US-09-850-799-1

```

	Query Match Best Local Similarity Matches 228;	12.4%; 74.0%;	Score 180; Pred. No. 41e-48;	DB 10;	Length 1752;	
	Conservative	0;	Mismatches 80;	Indels. 0;	Gaps 0;	
QY	879 AGCCGTCCTTCCAATATCCAGGGAGTGGGCGACATACAGCTATGAGCGAGTTCCTACTGGACTT	938				
Db	529 AGCCCTCGGGTACAGAAAGGACACGGGCGACATCCAGTTGTGGCAGTTTCTTACTGGAAGCTG	588				
QY	939 CTGTGGGACAGCTCCAACTCCAACTGCATCACCCTGGGAGGGGACACAAATGGGGAGTTTCAG	998				
Db	589 CTGGCAGACCGCCGGAAGCGCGGCTGCATCGCGTGGGAGGGGCGGCCAGCGGGAATTCAG	648				
QY	999 ATACACAGACCCCTGATGAAAGTGGCTCGGGCGTTGGGGAGAGAGGAAAGCAACCTAACATG	1058				
Db	649 CTCACCGACCCCGACGAGGTGGGCGACGGCTGGGGGCGAGGCGCAAGACCAAGCCCAATATG	708				
QY	1059 AACCTATGACAAACTGCAGCCGTGCACCTCGCTACTACTATGACAAACAAATATTATGACTTAA	1118				
Db	709 AACTAGACAGACGTTAAGTGCAGCATGCGTCTACTACTACGACAAAAACATCATGTAGCAAG	768				
QY	1119 GTTCATGTGTAACGCTATGCGCTTACAAATTTGATTTCCACGGAATGCGTCAAGGCGCCTCAG	1178				
Db	769 GTGCACGGGCAAGGCGTACGCGCTACCGCTTTGACTTCCAGGGGCGCTTGCAAGAGGCTTGCAG	828				
QY	1179 CCTCAGCC	1186				
Db	829 CCACCAAC	836				

```

RESULT 11
US-09-864-761-7427/c
; Sequence 7427, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

```

```

1  APPLICANT:  venk, Sharron G.
2  APPLICANT:  Rank, David R.
3  APPLICANT:  Hanzel, David K.
4  APPLICANT:  Chen, Wensheng
5  TITLE OF INVENTION:  HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
6  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
7  FILE REFERENCE:  Aecmca-x-1
8  CURRENT APPLICATION NUMBER:  US/09/864,761
9  CURRENT FILING DATE:  2001-05-23
10 PRIOR APPLICATION NUMBER:  US 60/180,312
11 PRIOR FILING DATE:  2000-02-04
12 PRIOR APPLICATION NUMBER:  US 60/207,456
13 PRIOR FILING DATE:  2000-05-26
14 PRIOR APPLICATION NUMBER:  US 09/632,366
15 PRIOR FILING DATE:  2000-08-03
16 PRIOR APPLICATION NUMBER:  GB 24263,6
17 PRIOR FILING DATE:  2000-10-04
18 PRIOR APPLICATION NUMBER:  US 60/236,359
19 PRIOR FILING DATE:  2000-09-27
20 PRIOR APPLICATION NUMBER:  PCT/US01/00666
21 PRIOR FILING DATE:  2001-01-30
22 PRIOR APPLICATION NUMBER:  PCT/US01/00667
23 PRIOR FILING DATE:  2001-01-30
24 PRIOR APPLICATION NUMBER:  PCT/US01/00664
25 PRIOR FILING DATE:  2001-01-30
26 PRIOR APPLICATION NUMBER:  PCT/US01/00669
27 PRIOR FILING DATE:  2001-01-30
28 PRIOR APPLICATION NUMBER:  PCT/US01/00665
29 PRIOR FILING DATE:  2001-01-30
30 PRIOR APPLICATION NUMBER:  PCT/US01/00668
31 PRIOR FILING DATE:  2001-01-30
32 PRIOR APPLICATION NUMBER:  PCT/US01/00663
33 PRIOR FILING DATE:  2001-01-30
34 PRIOR APPLICATION NUMBER:  PCT/US01/00662
35 PRIOR FILING DATE:  2001-01-30
36 PRIOR APPLICATION NUMBER:  PCT/US01/00661
37 PRIOR FILING DATE:  2001-01-30
38 PRIOR APPLICATION NUMBER:  PCT/US01/00670
39 PRIOR FILING DATE:  2001-01-30
40 PRIOR APPLICATION NUMBER:  US 60/234,687
41 PRIOR FILING DATE:  2000-09-21
42 PRIOR APPLICATION NUMBER:  US 09/608,408
43 PRIOR FILING DATE:  2000-06-30
44 PRIOR APPLICATION NUMBER:  US 09/774,203
45 PRIOR FILING DATE:  2001-01-29
46 NUMBER OF SEQ ID NOS: 49117
47 SOFTWARE:  Annonmax Sequence Listing Engine vers. 1.1
48 SEQ ID NO 7427
49 LENGTH: 553
50 TYPE:  DNA
51 ORGANISM:  Homo sapiens
52 FEATURE:
53 OTHER INFORMATION:  MAP TO AP001426.1
54 OTHER INFORMATION:  EXPRESSED IN PLACENTA, SIGNAL = 2.1
55 OTHER INFORMATION:  EXPRESSED IN BONE MARROW, SIGNAL = 2.1
56 OTHER INFORMATION:  EXPRESSED IN LUNG, SIGNAL = 1.3
57 OTHER INFORMATION:  EXPRESSED IN ADULT LIVER, SIGNAL = 0.76
58 OTHER INFORMATION:  EXPRESSED IN HEART, SIGNAL = 2.1
59 OTHER INFORMATION:  EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
60 OTHER INFORMATION:  EXPRESSED IN BRAIN, SIGNAL = 2
61 US-09-864-761-7427
62 Query Match 11.6%; Score 167.8; DB 10; Length 553;
63 Pred. Local Similarity 73.5%;
64 Matches 214; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

[illegible]


```

1  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.52
2  OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.45
3  OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.55
4  OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUATE 4.00e-38
5  OTHER INFORMATION: NT HIT: M21535.1, EVALUATE 1.00e-113
6  OTHER INFORMATION: EST_HUMAN HIT: AW948986.1, EVALUATE 1.00e-108
7  US-09-864-761-29364
8
9  Query Match 11.3%; Score 163.2; DB: 10; Length 348;
10 Best Local Similarity 86.5%; Pred. No. 5.2e-43;
11 Matches 180; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
12
13 448 CAGATCCTACGTTGTGGAGACAGACCATGTACGGCAGTGGCTGGAGTGGGAGTGAAG 507
14 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15 312 CAGATCCTACGCTATGGATGTACAGACCATGTGGCGCAGTGGCTGGAGTGGGCGTGAAG 253
16
17 508 AGTATGCTCTCCAGACGCTGCACATCTTGTGTTCTCCAGAACATGATGGGAAAGAGTTGT 567
18 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19 252 AATATGGCCTTCAGAGAGCTCAACATCTTGTATTCTCCAGAACATGTAGAGGAGCACTGT 193
20
21 568 GTAATAATGACCAAGATGATCTCCAGAGACTCCAGCGGAGGTATTAACGAGATATCTCTCC 627
22 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 192 GCAAGATGACCAAGAGAGACTCTCCAGAGGCTCCAGCCCAAGCTTACAGCGGAGACATCTTC 133
24
25 628 TGTCACACCTACCTACCTACCTCAGAGAGAG 655
26 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 132 TCTCACAATCTCCTACCTACCTCAGAGAGAG 105
28
29 RESULT 13
30 US-09-864-761-5035/c
31 Sequence 5035, Application US/09864761
32 Patent No. US20020048763A1
33 GENERAL INFORMATION:
34 APPLICANT: Penn, Sharron G.
35 APPLICANT: Rank, David R.
36 APPLICANT: Hanzel, David K.
37 APPLICANT: Chen, Wensheng
38 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
39 FILE REFERENCE: Aeomica-X-1
40 CURRENT APPLICATION NUMBER: US/09/864,761
41
42 CURRENT FILING DATE: 2001-05-23
43
44 PRIOR APPLICATION NUMBER: US 60/180,312
45
46 PRIOR FILING DATE: 2000-02-04
47
48 PRIOR APPLICATION NUMBER: US 60/207,456
49
50 PRIOR FILING DATE: 2000-05-26
51
52 PRIOR APPLICATION NUMBER: US 09/632,366
53
54 PRIOR FILING DATE: 2000-08-03
55
56 PRIOR APPLICATION NUMBER: GB 24263.6
57
58 PRIOR FILING DATE: 2000-10-04
59
60 PRIOR APPLICATION NUMBER: US 60/236,359
61
62 PRIOR FILING DATE: 2000-09-27
63
64 PRIOR APPLICATION NUMBER: PCT/US01/00666
65
66 PRIOR FILING DATE: 2001-01-30
67
68 PRIOR APPLICATION NUMBER: PCT/US01/00667
69
70 PRIOR FILING DATE: 2001-01-30
71
72 PRIOR APPLICATION NUMBER: PCT/US01/00664
73
74 PRIOR FILING DATE: 2001-01-30
75
76 PRIOR APPLICATION NUMBER: PCT/US01/00669
77
78 PRIOR FILING DATE: 2001-01-30
79
80 PRIOR APPLICATION NUMBER: PCT/US01/00665
81
82 PRIOR FILING DATE: 2001-01-30
83
84 PRIOR APPLICATION NUMBER: PCT/US01/00668
85
86 PRIOR FILING DATE: 2001-01-30
87
88 PRIOR APPLICATION NUMBER: PCT/US01/00663
89
90 PRIOR FILING DATE: 2001-01-30
91
92 PRIOR APPLICATION NUMBER: PCT/US01/00662
93
94 PRIOR FILING DATE: 2001-01-30
95
96 PRIOR APPLICATION NUMBER: PCT/US01/00661
97
98 PRIOR FILING DATE: 2001-01-30
99
100 PRIOR APPLICATION NUMBER: PCT/US01/00670
101
102 PRIOR FILING DATE: 2001-01-30

```

RESULT 15
 US-09-864-761-21761/c
 : Sequence 21761, Application US/09864761
 Patent No. US20020048763A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Penn, Sharon G.
 : APPLICANT: Rank, David R.
 : APPLICANT: Hanzel, David K.
 : APPLICANT: Chen, Wensheng
 :
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 : TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 :
 : FILE REFERENCE: Meomica-X-1
 :
 : CURRENT APPLICATION NUMBER: US/09/864,761
 :
 : CURRENT FILING DATE: 2001-05-23
 :
 : PRIOR APPLICATION NUMBER: US 60/180,312
 :
 : PRIOR FILING DATE: 2000-02-04
 :
 : PRIOR APPLICATION NUMBER: US 60/207,456
 :
 : PRIOR FILING DATE: 2000-05-26
 :
 : PRIOR APPLICATION NUMBER: US 09/632,366
 :
 : PRIOR FILING DATE: 2000-08-03
 :
 : PRIOR APPLICATION NUMBER: GB 24263.6
 :
 : PRIOR FILING DATE: 2000-10-04
 :
 : PRIOR APPLICATION NUMBER: US 60/236,359

```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21761
; LENGTH: 225
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000021.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.95
; OTHER INFORMATION: NT HIT: M21535.1, EVALUE 1.00e-113
; OTHER INFORMATION: EST HUMAN HIT: AW948986.1, EVALUE 1.00e-108
; OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 3.00e-37
US-09-864-761-21761
```

```

Query Match      11.2%; Score 162.2; DB 10; Length 225;
Best Local Similarity 86.5%; Pred. No. 8.4e-43;
Matches 179; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY      448 CAGATCTACGTAATGAGCAGACAGACATGACGAGTGCGAGTGGGCACTGAAG 507
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      207 CAGATCTACGTAATGAGCAGACATGACGAGTGCGAGTGGGCGGTGAAG 148

QY      508 AGTATGCTCTCCAGAGTGACATCTGTGTTCCAGAACATGATGGGAAGATTGT 567
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      147 AATATGCTCTCCAGAGTGACATCTGTGTTCCAGAACATGATGGGAAGAACTGT 88

QY      568 GTAATATGACCAAGATGACTCCAGAGACTACGCGGAGCTATATACGAGATATCTCC 627
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      87 GCAAGATGACCAAGAGAGACTCTCCAGAGGCTCACCCGAGCTACAGCCGACATCTCTC 28

QY      628 TGTCACTACTACTACTACTAGAGAGA 654
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      27 TCTCAGATCTCTCACTACTACTAGAGAGA 1
```

Search completed: July 28, 2003, 12:23:28
Job time : 400.967 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 09:24:34 ; Search time 2774.91 Seconds
(without alignments)
8445.286 Million cell updates/sec

Title: US-09-902-772-1

Perfect score: 1447
Sequence: 1 gaattccgcgaacgaatrat.....gaaagaagaagcccaagaana 1447

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_frod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485.4	33.5	880	14	BQ233264 AGENCOURT
2	482.2	33.3	715	12	BG388025 602412867
3	475.4	32.9	642	13	BT173036 1601408.Y
4	468.6	32.4	701	9	AU136709 AU136709
5	468.6	32.4	1089	14	BQ212127 AGENCOURT
6	462	31.9	634	13	BM489636 pgm2n.pk0

7	452.2	31.3	668	10	BB660034
8	445.2	30.8	915	12	BG390291
9	444.2	30.7	928	12	BG259459
10	442	30.5	1045	13	BM456833
11	436.2	30.1	675	10	BB639043
12	432.6	29.9	629	10	BB577887
13	428.8	29.6	629	10	BM448986
14	425.6	29.4	592	13	BM088063
15	380	26.3	827	9	AJ456498
16	367.8	25.4	522	12	BF566705
17	367.2	25.4	516	12	BG691566
18	352.8	24.4	539	13	BT558012
19	335	23.2	553	12	BG256864
20	317.2	21.9	671	13	BT558530
21	313.4	21.7	740	9	AJ456506
22	313	21.6	1074	14	BQ953977
23	308	21.3	666	12	BG390572
24	304.8	21.1	645	10	BB656030
25	298.2	20.6	651	10	BM442803
26	297.8	20.6	1245	14	BQ212133
27	297.4	20.6	524	9	AJ794504
28	293	20.2	420	13	BT181694
29	284.6	19.7	674	10	BB238163
30	261.6	18.1	1015	9	AL554108
31	259.8	18.0	400	10	BE245360
32	256.2	17.7	619	10	AM421310
33	241	16.7	512	12	BE831331
34	232.6	16.1	327	10	AM209658
35	232.4	16.1	634	10	BB655125
36	228.6	15.8	465	14	R87572
37	218.8	15.1	551	9	AA467723
38	215.8	14.9	815	9	AJ446143
39	214.2	14.8	1068	14	BQ062565
40	211	14.6	520	10	BE233460
41	209	14.4	756	9	AJ445352
42	208	14.4	872	14	BQ232274
43	201.4	13.9	788	9	AJ456869
44	200.2	13.8	675	13	BM440515
45	198	13.7	874	14	BQ215477

ALIGNMENTS

RESULT 1
BQ233264
LOCUS
DEFINITION AGENCOURT.7565836 NIH_MGC_92 Homo sapiens CDNA IMAGE:6041526
5' mRNA sequence.
ACCESSION BQ233264
VERSION BQ233264
KEYWORDS BQ233264.1 GI:20414664
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 880)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMU at:
http://image.llnl.gov
Plate: L1AM13279 row: h column: 07
High quality sequence stop: 608.
Location/Qualifiers
1. .880

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6041526"
/clone_lib="NIH MGC 92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC library."

```

BASE COUNT	250 a	231 c	207 g	189 t	3 others
ORIGIN					

Query Match	33.5%;	Score 485.4;	DB 14;	Length 880;
Best Local Similarity	84.4%;	Pred. No. 9e-139;		
Matches 546;	Conservative 0;	Mismatches 101;	Indels 0;	Gaps 0;

OY	8	GCGAAGCAATATATTTTATTTTACCAATTTTATACGATCAATATATCTGTACATATATGGC	67
Db	19	GAGAAAGCAATATTTTATTTATTAACATTTTATTAACGATCAATATAACTGTATTCATTTATGGC	78
OY	68	AAGCACTATTAAGGAAGCAATTATCATGCGGTGAGTGAAGACAGCTCCCTTTTGAGTGTGC	127
Db	79	CAGCACTATTAAGGAAGCAAGCTTATCATGTTGTGAGTGGAGACAGTCCGTTGTGAGTGTGC	138
OY	128	CTAGCGATCGCCCACTTGTGCAGAAACAGAAATGACAGCTCTCTCTTCCAGTGAATATGG	187
Db	139	CTACGGAAAGCCACACCTGGCTGTAAACACAGAGATGACCGGCTCTCTCCAGGAGCATATGG	198
OY	188	GCAAACATCAAAAGATGACCCGCGGGTTCGCCAGAGAGACTGGTTATCAACACCCCCGGC	247
Db	199	ACAGACTCCAAAGATGAGCCCAACGGCTGCCCTGACGAGATGGGCTGTCTCAACCCCAAC	258
OY	248	CAGAGTTACCACTTAAGATGAGGTGTATACCCAAACAGATTAATGGGTCAAGAAATTCACC	307
Db	259	CAGGCTACCACTCAAAATGGAATGTATACCTTCGCCAGTGAATGGTCTCAAGGAACCTCTCC	318
OY	308	TGATGACACGACGCGGCGCAAAAGGAGGGAATAATGGCTTACACACTTCAGAAATGTTGGGAT	367
Db	319	TGATGAAATGCAAGTGTGGGCCAAAGGGGGGAAGATGGTGGGACACCCAGACACCGTTGGGAT	378
OY	368	GAACTATGGAAGCTACATGAGAGAGAGAGACATATTCGCGCTCCAAATATGACAAACCAATGA	427
Db	379	GAACTACGCGACGCTACATGAGAGAGAAACACATCTGCACCCCCCAACATGACCAACGAAGA	438
OY	428	ACGAAGAATATTTGTGCCAGCAGATCCCTACGTTATGAGACACAGACACATGTACGGAGTGG	487
Db	439	GCGCAGATATTTTCGTGCCAGCAGATCCCTACGTTATGAGATTAACAGACCAATGTCCGGCAGTG	498
OY	488	GCTGGAGTGGGACAGTGAAGAGAGTATGTCTTCCACAGCGGACGACATCTTTGTCCAGAA	547
Db	499	GCTGGAGTGGGCGGTGAAGAAATATATAGGCTTCCACAGCGTCAACACATCTTTTATTCAGAA	558
OY	548	CATTGATGGGAAAAGATGTGTAAATGACCAAGAATGACTTCCAGAGACTACGCGCGAG	607
Db	559	CATGATGGGGAAGGAACGTGCACAAATGACCAAGAGACACTTCCAGAGGCTCACCCCCAG	618
OY	608	CTATAAGCAATATCTCTCTGTCACACCTACACTACCTACCAAGAGA	654
Db	619	CTACACGCGGACACCTCTTCTCTACACTCTCCACTACTCTCAGAGAGA	665

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 715)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

1. [cDNA Library Preparation](#): Life Technologies, Inc.
 2. [cDNA Library Arrayed by the I.M.A.G.E. Consortium \(LLNL\)](#)
 3. [DNA Sequencing by: Incyte genomics, Inc.](#)
 4. [Clone distribution: MGC clone distribution information can be](#)
 5. [found through the I.M.A.G.E. Consortium/LLNL at:](#)
 6. <http://image.llnl.gov>
 7. [Plate: LLNL0420 row: f column: 18](#)
 8. [high quality sequence stop: 693.](#)

FEATURES	Location/Qualifiers
source	1. .715

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:452123"
/clone_1db="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies Note: this is a NIH_MGC library."

```

BASE COUNT	206 a	188 c	172 g	149 t
ORIGIN				

Query Match	33.3%	Score 482.2;	DB 12;	Length 715;
Best Local Similarity	84.1%;	Pred. No. 8e-138;		
Matches 544;	Conservative 0;	Mismatches 103;	Indels 0;	Gaps 0;

[illegible]

Query Match	32.9%;	Score 475.4;	DB 13;	Length 642;
Best Local Similarity	84.0%;	Pred. No. 9.7e-136;		
Matches 536;	Conservative 0;	Mismatches 102;	Indels 0;	Gaps 0;
QY	15	AATTAATATTTATGCAATTTATAGCATCATTAATCTTGATCATATTATGGCAAGCACT	74	
	4	AATTAATATTTATTAACATTTATTAACCATCATTAATAACTTGATTCGCAATTATGGCCAGCACT	63	

Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project: 5' - & 3' -end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 Location/Qualifiers
 1..701

Db 621 GTCATGCGGACATCTCTCTCCATCTCCTACTACTAGAGAGA 668

RESULT 6
BM489636
LOCUS
DEFINITION
634 bp mRNA linear EST 07-FEB-2002
pkm2n.pk011.124 Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal growth plate cDNA library (pkm2n) Gallus gallus CDNA
TRANSCRIPTIONAL REGULATOR ERG p1560754 transcription factor erg
chicken emb|CA54404.1| (x77159) ERG [Gallus gallus], mRNA
sequence.
BM489636
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Chicken.
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 634).
Cogburn, L.A. and Monsonego-Ornan, E.
ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate CDNA Library, USDA/IRAFS Animal Genome
Project
Unpublished (2002)
JOURNAL
COMMENT
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES
Source
Location/Qualifiers
1..634
/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
strains 90 & 21"
/db_xref="taxon:9031"
/clone="pkm2n.pk011.124"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate CDNA Library (pkm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast, leg: Embryo(d19); post-hatch(1d,1,3,5,7,9
11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMDH10B"
/note="Vector: pcMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"

BASE COUNT
192 a 164 c 138 g 140 t

ORIGIN
Query Match 31.9%; Score 462; DB 13; Length 634;
Best Local Similarity 87.2%; Pred. No. 1.4e-13;
Matches 553; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 418 CAACCAATGAAGAGATTATGTCAGCAGATCTTACGTTATGAGCAGACCATG 477
Db 1 CAACCAATGAAGAGATTATGTCAGCAGATCTTACGTTATGAGCAGACCATG 60
QY 478 TAGCGCATGCGTGGAGTGGGAGTGAAGGATGTCCTTCAGAGCTGGACATCTTGT 537
Db 61 TAGCGCATGCGTGGAGTGGGAGTGAAGGATGTCCTTCAGAGCTGGACATCTTGT 120
QY 538 TGTTCAGACATGATGAGAAAGTGTGTAAATGACCAAGATGACTTCCAGAGAC 597
Db 121 TGTTCAGACATGATGAGAAAGTGTGTAAATGACCAAGATGACTTCCAGAGAC 180
QY 598 TCACGCCAGGTATTAACGACATATCTCTGTCCACACTACTACTCAGAGAGA--- 654

Db 181 TCAGCCGACGATATTAACGAGATATCTCTCTGTCACACTACTACTCAGAGAGATC 240
QY 655 ----- 654
Db 241 CTCCTCCACATTTGACCTTCAGATGATTTGATTAAGCCCTTACAAAACCTCCACGTTAA 300
QY 655 -----GAGAGCCACTTTTATTTTCCAAATCATCATGTTTACCAG 696
Db 301 TGCATCTAGAAACACAGAGAGAGGACTTTTATTTTCCAAATCATCATGTTTACCAG 360
QY 697 AAGCAACCCAAAGATTAACAACAAGGCCAGATTTACCTTATGAGCAAGCGAGAGATCAG 756
Db 361 AAGCAACCCAAAGATTAACAACAAGGCCAGATTTACCTTATGAGCAAGCGAGAGATCAG 420
QY 757 CGTGCAGACATTCACAGCCATCCCTACGTAGTAAAAAGCTACCCACATCATCTTCAACAG 816
Db 421 CGTGCAGACATTCACAGCCATCCCTACGTAGTAAAAAGCTACCCACATCATCTTCAACAG 480
QY 817 TGCCCAAAACAGAAAGACAGCGCTCCTAGTATGATCCTTATCAGATTTCTTGAGCCAGCA 876
Db 481 TGCCCAAAACAGAAAGACAGCGCTCCTAGTATGATCCTTATCAGATTTCTTGAGCCAGCA 540
QY 877 GCAGCCGCTTTCGAATTCAGAGAGAGTGGGAGATACAGTATGCGAGTTCTACTGAGC 936
Db 541 GCAGCCGCTTTCGAATTCAGAGAGAGTGGGAGATACAGTATGCGAGTTCTACTGAGC 600
QY 937 TTCTGTGCGACAGCTCCACATCTCCACATGATCAG 970
Db 601 TTCTGTGCGACAGCTCCACATCTCCACATGATCAG 634

RESULT 7
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
TITLE
JOURNAL
COMMENT

BB660034 668 bp mRNA linear EST 26-OCT-2001
BB660034 RIKEN full-length enriched, 13 days embryo lung Mus
musculus CDNA clone DA30027104 5', mRNA sequence.
BB660034
BB660034.1 GI:16493855
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K., Itoh,
M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E.,
Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuiura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format

	BASE COUNT	265 a	242 c	220 g	201 t	
ORIGIN						
Query Match	30.7%;	Score 444.2;	DB 12;	Length 928;		
Best Local Similarity	83.5%;	Pred. No. 5.5e-126;				
Matches 540;	Conservative	0;	Mismatches 103;	Indels 4;	Gaps 3;	
Oy	8	GCGAACCAATATTATTATTTAGCAATTAATGACAAATAAATCTGCATCATATTAGC	67			/note="Organ: testis; Vector: PCMV-SPOB6; Site_1: NCBI; Site_2: Sall; Cloned unidirectionally; oligo-dt primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: This is a NIH-MGC library."
Dd	33	GAGAGAACAATATTATTATTATTAACATTTTATTAACATCAATAAATCAATTGGCATTAGCC	92			
Oy	68	AAGCACTATTAAAGNACATTAATCATGTGAGTGGAAGACCAGTCCTTGTTTGAGTGTGC	127			
Dd	93	CAGACACTATTAAAGAACGCTTATAGTGTGTAGTGAGACACAGT-CGTGTTTGAGTGTGC	151			
Oy	128	CTACGGATTCGCCCACCTTGCCAAGACAGAAAATGACAGCCTCTCTTCAGTGAATATGG	187			
Dd	152	CTACGGAAAGCGCAACCTGGCTTAGACAGAGATGACCGCGCTCCCTCCAGCGCACTATGG	211			
Oy	188	GCAAACATCAAAAGATGAGCCCCGCCCGCTTCCCACAGAGACTGCTTATCACAGCCCCGCCG	247			
Dd	212	ACAACACTTCCAAAGATGAGCCCGCCGCTCCACAGAGATGCGTGTCTCAACCCCGAGC	271			
Oy	248	CAGAGTTACCTTTAAGATGAGTGTAAACCCAAACAGAGTTAATGGGTCAGAGAAATTCACC	307			
Dd	272	CAG-GTCACTCATCAAAATGMAATTTAAACCTCAGCGAGTGAATGCTCAAGAAATCTCC	330			
Oy	308	TGATGACTGCAGCGTGGCAAAAGAGGAAAAATGGTTAGCAGTTCAACAATGTTGGGAT	367			
Dd	331	TGATGAATGCGAGTGTGGCCAAAGGCGGGAAAGATGGTGGGAGAGCCCAACACCTTGGGAT	390			
Oy	368	GAAGTATGGAAGCTACATGGAAGAAGACATATGCCCTCCAAATATGACAACCATATGA	427			
Dd	391	GAACTACGGCAGCTPACATGGAAGAGAAACACATGCCACCCCCAAACATGACCAAGAAAGA	450			
Oy	428	ACGAGAAGTTTATGTGCAGACAGATCCTACGTTATGGAGACACAGACCATGATAGCGCAGTG	487			
Dd	451	GCGCAGAGATTATCTGCCAGACAGATCCTACGCTTGGAGATACAGACCATGTGGCGCAGTG	510			
Oy	488	GCTGAGTGGGCACTGAGGAAGATGTGCTTCCAGACGTGACATCTTGTGTGTTGTCAGAA	547			
Dd	511	GCTGAGTGGGCGGTGACAGAAATATGGCTTCCAGACGTCAACATCTTGTTATTCAGAA	570			
Oy	548	CATTGATGGGAAGAAGTTGTCTAAATGACCAAAAGATGACTTCCAGAGACTACGCCGAG	607			
Dd	571	CATGATGAGGGGAAGACCTGTGCAGAGATGACCAAGACGACTTCCAGAGGCGTCACCCC--A	628			
Oy	608	CTAATAGCGAGATATCTCTCGTCAGACCTAACATACCTCAGAGAGA	654			
Dd	629	GTACAAATGCGACATCTCTCTCACATCTCCACTACCTCAGAGAGA	675			
RESULT 10						
LOCUS	BM456833	1045 bp	mRNA	linear	EST 05-FEB-2002	
DEFINITION	AGNCUCURT_6404041 NIH_MGC_92 Homo sapiens CDNA IMAGE:5383753					
VERSION	BM456833					
KEYWORDS	BM456833.1 GI:18505873					
SOURCE	EST.					
ORGANISM	human.					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
REFERENCE	1 (bases 1 to 1045)					
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.					
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: c9apds-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM12347 row: f column: 10
 High quality sequence stop: 669.
 Location/Qualifiers

FEATURES

source

1. 1045
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5583753"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 300 a 290 c 230 g 225 t
 ORIGIN

Query Match

Best Local Similarity 76.6%; Score 442; DB 13; Length 1045;
 Pred. No. 2.8e-125; Mismatches 165; Indels 14; Gaps 3;

Matches 585; Conservative 0; Mismatches 165; Indels 14; Gaps 3;
 81 GAAGCATATTCAGTGTGAGTGAAGACAGTCTTGTGTAGTGTGCTTACGATCCG 140
 190 GAAGCCTTATCAGTGTGAGTGAAGACAGTCTTGTGTAGTGTGCTTACGATCCG 249
 141 CACCTGCAAGAGCAAGTGAAGACAGTCTTGTGTAGTGTGCTTACGATCCG 200
 250 CACCTGCAAGAGCAAGTGAAGACAGTCTTGTGTAGTGTGCTTACGATCCG 309
 201 ATGAGCCCGCGCTTCCCGACAGAGTGTATACAGCCCGCGCGCAAGTACCAT 260
 310 ATGAGCCCGCGCTTCCCGACAGAGTGTATACAGCCCGCGCGCAAGTACCAT 369
 261 AAGATGAGTGAATCCCAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 320
 370 AAGATGAGTGAATCCCAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 429
 321 GTGGCAAGAGGAGGAAATGTTAGATTCAGATTCAGATTCAGATTCAGATTCAG 380
 430 GTGGCAAGAGGAGGAAATGTTAGATTCAGATTCAGATTCAGATTCAGATTCAG 489
 381 TACATGAGAGAGAGATATTCGCGCTCCCAATATATGACCAATGAGAGAGATAT 440
 490 TACATGAGAGAGAGATATTCGCGCTCCCAATATATGACCAATGAGAGAGATAT 549
 441 GTGCGACAGATCTTACGTTATGAGACAGACAGATTCAGAGAGTGGCGTGA 500
 550 GTGCGACAGATCTTACGTTATGAGACAGACAGATTCAGAGAGTGGCGTGA 609
 501 GTGAGAGATGAGTGTCTTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 560
 610 GTGAGAGATGAGTGTCTTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 669
 561 GAGTGTGTAAATGAGCAAGATGAGTTCAGAGAGTGAAGTGAAGTGAAGTGAAG 620
 670 GAGTGTGTAAATGAGCAAGATGAGTTCAGAGAGTGAAGTGAAGTGAAGTGAAG 729
 621 ATCTCTCTGCAACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC 680
 730 ATCTCTCTGCAACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC 789
 681 ACATGAGTTTACCAAGAGCAAC-----GCAAGAAATATACACAGCCGATTTAC 732

Db

790 GATGTGATTAAGGCTTACCAACTCTCCCGGTTAATGATGCTAGAACACCGCATTTAC 849

Qy

733 CTATATGAGCAAGGAGAGATGACCGTGGAGAGTCAAGCAGATCC-----ACTGAGTCA 788

Db

850 CATRTGAGACCCCCCGGAGATGACCTTGACCGGTGACGCTCAACCCCAAGCCCAAGTCG 909

Qy

789 AAGGCTACCAACCATCAT--CTTCAACAGTGGCCCAACAGAA 830

Db

910 AAGGCTTCAACATCTCCCTTCAACAGAGCCCAAACTGAA 953

RESULT 11

BB639043

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB639043 675 bp mRNA linear EST 31-AUG-2001
 BB639043 RIKEN full-length enriched, 3 days neonate thymus Mus
 musculus cDNA A630068021 5', mRNA sequence.
 BB639043 1 GI:15401323
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 675)
 Arekawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arekawa, T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshinori Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
 M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura,
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
 Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Yamana, K., Kiyosawa, H., Kondo, S., Saito, T., Shingawa, A., Aizawa,
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T.,
 Ishii, Y., and Hayashizaki, Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
 Func. Genomics 2 pre, 172-186 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1. 675
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A630068021"
 /clone_lib="RIKEN full-length enriched, 3 days neonate

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:14:59 ; Search time 3594.89 Seconds
(without alignments)
3651.115 Million cell updates/sec

Title: US-09-902-772-2
Perfect score: 2444
Sequence: 1 MASTIKALSVSEDSLFE.....IYENTRLPAHNPMSHIGTY 451

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ -p2n.model -DEV-xlh
-O/cgn2_1/USPTO.spool/US09902772/runat.23072003_093657_14878/app_query.fasta.1.1294
-DB-GenEmbl -GFMT-fastap -SUFFIX-rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-Dlosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09902772_6CGN_1_1_4707_etrnat.23072003_093657_14878 -NCPU=6 -ICPU=3
-NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_da.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444	100.0	1413	5 AY065661	AY065661 Gallus ga
2	2444	100.0	1447	6 E31253	E31253 Protein hav
3	2414.5	98.8	1516	5 GGERG	X77159 G.gallus ER
4	2414.5	98.8	1528	6 E31254	E31254 Protein hav
5	2267	92.8	2209	10 AB073080	AB073080 Mus muscu
6	2192	89.7	2012	5 XLAJ4126	AJ224126 Xenopus l
7	2182	89.3	3166	9 HUMERG2	M17254 Human erg2
8	2179	89.2	1808	10 AB031088	AB031088 Rattus no
9	2159.5	88.4	2133	10 AB073079	AB073079 Mus muscu
10	2153.5	88.1	1798	5 XLAJ4125	AJ224125 Xenopus l
11	2151	88.0	2131	10 AB073078	AB073078 Mus muscu
12	1810	74.1	3126	9 HUMERG11	M21535 Human erg p
13	1637	67.0	1436	5 DRE249590	AJ249590 Danio rer
14	1624.5	66.5	1359	9 AY029368	AY029368 Homo sapi
15	1624.5	66.5	2394	9 BC010115	BC010115 Homo sapi
16	1624.5	66.5	2916	6 A36461	A36461 Sequence 2
17	1624.5	66.5	2938	6 AK080101	AK080101 Sequence
18	1624.5	66.5	2938	9 HSHWFLI	X67001 R.septens H
19	1624.5	66.5	2957	9 HUMERGBFL1	M88833 Homo sapien
20	1624.5	66.5	2959	9 BC001670	BC001670 Homo sapi
21	1624.5	66.5	6849	12 AY029367	AY029367 Synthetic
22	1621.5	66.3	1673	9 S45205	S45205 Flt-1-Frien
23	1613.5	66.0	1729	10 MMFLI1	X59421 Mouse Flt-1
24	1579.5	64.6	3490	5 CCRNAFLI	Y14773 Coturnix co
25	1576	64.5	4403	5 XLFILG	X66979 X.laevis MR
26	1567.5	64.1	3545	5 CCFILONCO	Y14774 Coturnix co
27	1536.5	62.9	1932	9 HUMERFLIA	M83255 Human Flt-1
28	1486.5	60.8	1245	5 AF177538	AF177538 Danio rer
29	1227	50.2	816	10 S66169	S66169 Erg-3-Immum
30	1183.5	48.4	762	9 S72621	S72621 EMS...erg 1
31	965	39.5	1431	9 AF327066	AF327066 Homo sapi
32	956.5	39.1	14518	9 AP001423	AP001423 Homo sapi
33	956.5	39.1	100000	9 AP000021	AP000021 Homo sapi
34	956.5	39.1	100000	9 AP000163	AP000163 Homo sapi
35	956.5	39.1	340000	9 AP001731	AP001731 Homo sapi
36	925	37.8	588	9 S72865	S72865 EMS...EMS-e
37	833.5	34.1	66352	9 HSY17293	Y17293 Homo sapien
38	833.5	34.1	187131	9 AP001535	AP001535 Homo sapi
39	833.5	34.1	217382	9 AP001122	AP001122 Homo sapi
40	682	27.9	634	3 SULERG	M61067 Sea urchin
41	664	27.2	2180	3 AY060316	AY060316 Drosophila
42	581.5	23.8	491	3 HDJ311813	AJ311813 Helicoste d
43	543.5	22.2	482	3 PCU278692	PCU278692 Perinereid
44	520.5	21.3	169741	9 AC097468	AC097468 Homo sapi
45	519.5	21.3	192968	2 AC127107	AC127107 Rattus no

RESULT 1

ALIGNMENTS

AY065661
LOCUS AY065661 1413 bp mRNA linear VRT 17-DEC-2001
DEFINITION Gallus gallus erg isoform C-1-1 mRNA, complete cds; alternatively spliced.
ACCESSION AY065661
VERSION AY065661.1 GI:17887440
KEYWORDS
SOURCE
ORGANISM Gallus gallus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE
AUTHORS 1 (bases 1 to 1413)
Iwamoto, M., Higuchi, Y., Koyama, E., Enomoto-Iwamoto, M., Kurisu, K., Yeh, H., Abrams, W.R., Rosenbloom, J. and Pacifici, M.
TITLE Transcription factor ERG variants and functional diversification of chondrocytes during limb long bone development
J. Cell Biol. 150 (1), 27-40 (2000)
JOURNAL MEDLINE 20351415
PUBMED 10893354
REFERENCE
AUTHORS 2 (bases 1 to 1413)
Iwamoto, M., Higuchi, Y., Enomoto-Iwamoto, M., Kurisu, K., Koyama, E., Yeh, H., Rosenbloom, J. and Pacifici, M.
TITLE The role of ERG (ets related gene) in cartilage development
Osteoarthritis. Cartil. 9 Suppl A, S41-S47 (2001)
JOURNAL MEDLINE 21535378
PUBMED 11680687
REFERENCE
AUTHORS 3 (bases 1 to 1413)
Iwamoto, M., Higuchi, Y., Yeh, H. and Pacifici, M.
TITLE Direct Submission
Submitted (05-DEC-2001) Oral Anatomy and Developmental Biology, Osaka University Faculty of Dentistry, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan
FEATURES
source location/Qualifiers
1..1413
/organism="Gallus gallus"
/db_xref="taxon:9031"
37..1392
/note="alternatively spliced; transcription factor"
/codon_start=1
/product="erg isoform C-1-1"
/protein_id="FAL40889.1"
/db_xref="GI:17887441"
/translation="MASTIKREALSVSESDSLFECAYGSPHLAKTEKMTSSSEYQOT SKSPVPQDDWLSOPPARVTIMECNPQVNSRNSPDCSAVAGKMYSSSDNYGM NGSYMEKRIIPPMNTTNERVIVPADPTLMSTDHVRQLEWAVEXGLPDYDLFE QNIDKEELCKMTKODFORLTPSYNADILSLHLYREGRDFFEPMTSVYEPATORIT TRPDLPYEOARSAMTSHSPHOSKATOPSSSVPTEDORPOLDPOYLIGPSSRLA NPSSGOLIMOFLELLISDSNSNCITMECTNECFKTDDEYARRKGEKSPNNAY DKLSRALRYTYDKNINMTKYHGRKRYATKFDHGIQALDOPHPESMTKIPSDLPYMS YHNPQKMNENVADHPALPVTSSSFPAAPNPYNSPTGIIYPTRLPAAHMPSHLGTYY"
BASE COUNT 424 a 370 c 311 g 308 t
ORIGIN
Alignment Scores:
Pred. No.: 1,4e-161 Length: 1413
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0
US-09-902-772-2 (1-451) x AY065661 (1-1413)

QY 1 MetAlaSerThrIleuysGluAlaIleuSerValIValSerGluAspGlnSerLeuPheGlu 20
Db 37 ATGGCAAGCACTTTAAGAGCAITATACGTGTGTGAGTGAAGACCAAGTCTTGTGTAG 96
QY 21 CysAlaTyrGlySerProHisLeuAlaIAluTyrGluMetThrAlaSerSerSerGlu 40
Db 97 TGTGCGTACGAGTGGCCCACTTGCAGAGACAGAAATGACACGCTCCTTCACAGGAA 156

QY 41 TyrGlyIleThrSerLysMetSerProArgValProGlnIleAspTyrLeuSerGlnPro 60
Db 157 TATGGCAAGCAATCAAAAGATGAGCCCGCGGTTCGCCAGCAGAGATGTTATCACAGCCC 216
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnIleValAsnGlySerArgAsn 80
Db 217 CCGGCGAGAGTTACCATTAAGATGAGGTGAACCCAAACAGGTTAATGGGTGAAGCAAT 276
QY 81 SerProAspPheCysSerValAlaIAluysGlyGlyLysMetValSerSerSerAspAsnVal 100
Db 277 TCACCTGATACCTGACGACCGTGGCAAAAGAGGAAATGGTTAGCAGTTGACATCANTGT 336
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
Db 337 GGGATGACATGATGGAAGCTACATGGAAGAAACACATATTCGCGCTCAATATGACAACC 396
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
Db 397 AATGAACGAAGAGTTATGTGCCAGCAGATCTTACGTTATGAGACACAGCCATGTGACG 456
QY 141 GlnTyrLeuGluTyrPAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
Db 457 CAGTGGCTGAGAGGGCAGTGAAGAGTATGCTTCCAGACCTGGACATCTTGTGTTC 516
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 517 CAGAACATGATGGAAGAGTGTGTAAATGACCAAGATGACTTCCAGACATCCAGC 576
QY 181 ProSerTyrAsnIleAspIleLeuSerHisIleHisTyrLeuArgGluArgValAla 200
Db 577 CCGAGCTATACCGAGATATCTCTGCTCACCTTACCTCAGAGAGAGAGAGAGCC 636
QY 201 ThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArg 220
Db 637 ACTTTTATTTTCCAAATATATCATGTTTACCAGGAAGCAACGAAAGATATCAACAAG 696
QY 221 ProAspLeuProTyrGluGlnAlaIAluArgArgSerAlaTyrThrSerHisSerHisProThr 240
Db 697 CCGAGTTTACCTTATGAGCAAGCAGAGAGATGACGTGCGAGTCAACGCCATCCCACT 756
QY 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260
Db 757 CAGTCAAAAGCTTACCAACCATATCTTCAACAGTGCACCAAAACAGAACCCAGCTCCT 816
QY 261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerArgLeuAlaAsnProGlySer 280
Db 817 CAGTTAGATCCTATACAGATTTCTTGACAGCAGACAGCCGTTCTCAATATCAGGAGCT 876
QY 281 GlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnAsn 300
Db 877 GGGCAAGATACAGTATGAGCAGTTCTTACGTGAGACTTCTGTGGACAGCTTCAACTCAAC 936
QY 301 CysIleThrTyrPLeuGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320
Db 937 TGCATACCTCGGAGGACACAAATGGGAGCTTCAATGACACAGACCTTATGACAGTGCCT 996
QY 321 ArgArgTyrGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340
Db 997 CGGCGTGGGAGAGAGAAAGCAACCTAACATGACATGACAAACCTCAGCGCTGCA 1056
QY 341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360
Db 1057 CTTGCGCTACTACTATGACAAAATATATGACTTAAGTTATGATGTAAGCGTATGCTTAC 1116
QY 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGluSerSerMet 380
Db 1117 AATTTGATTTTCCAGGAATCGCTCAGCGCTTCCACCTCACCCTCACAATATCATTCAAG 1176
QY 381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400
Db 1177 TACAATATACCATCAGACCTCCCTTACATGAGTCTTACATGACACACCCCAAGAGATG 1236
QY 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420

|||||
Db 1237 AACCTTGAGCTCCCGCATCCCGCTTGGCCCGGAACCTCATCCAGCTTTTGGCTGCC 1296
Qy 421 ProAsnProTyrrTPAsnSerProthrgLysLylleTyrrProAsnThrArgLeuProAla 440
Db 1297 CCTATCATCATCTGGAATTCACCAACTGAGCATCTACCCCAATATACAGAGCTGCACACT 1356
Qy 441 AlaHisMetProSerHisLeuGlyThrTyrr 451
Db 1357 GCTCATATGCTTCCTCCATCTTGGCACTTACTAC 1389
RESULT 2
E31253
LOCUS E31253
DEFINITION Protein having cell calcifying inhibitory activity and gene
encoding the same.
E31253
ACCESSION E31253.1 GI:13025685
VERSION JP 1999075871-A/1.
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1447)
AUTHORS Hiroyasu, I., Yoshinobu, H., Marjio, P., Joel, R. and Helena, E.
TITLE Protein having cell calcifying inhibitory activity and gene
encoding the same
JOURNAL Patent: JP 1999075871-A 1 23-MAR-1999;
CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF
PENNSYLVANIA
COMMENT OS Unidentified
PN JP 1999075871-A/1
PD 23-MAR-1999
PF 29-MAY-1998 JP 1998166076
PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI
HIGOYASU IWAMOTO, YOSHINOBU HIGUCHI, MARJIO PASHIFIKI, PI JOEL
ROZENBROOM,
PI HELENA E
PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00,
PC C12P21/02, C12N15/00, A61K37/02
PC C12P21/08, C12N15/00, A61K37/02
CC Strandedness: Single;
CC Topology: linear;
FH Key Location/Qualifiers
FT source 1..1447
FEATURES
source Location/Qualifiers
1..1447
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 440 a 374 c 317 g 316 t
ORIGIN
Alignment Scores:
Pred. No.: 1,44e-161 Length: 1447
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-902-772-2 (1-451) x E31253 (1-1447)
Qy 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 63 ATGGCAACACTATTAAAGAAAGCATTAACAGTGGAGTGAAGCACTCTGTTGTAG 122
Qy 21 CysAlaTyrrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
Db 123 TGTGCTTACGATGCGCCCACTTGCACAAAGACAGAAATGACAGCTCTCTTCCAGTGA 182
Qy 41 TyrrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrrLeuSerGlnPro 60
Db 183 TATGGGCAACATCAAAAGATGAGCGCGGCTTCCCGCAGCAAGCACTGTTATCAAGCCC 242

Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 243 CCGGCCAGAGTTTACCATTAAGATGAGTGAACCCAAACCGTTAAAGGTCAAGGAT 302
Qy 81 SerProAspPArgSerSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 303 TCACCTGATGACTGACGCGTGGCAAAAGAGGAGGAAATGTTAGACAGTTGACAACTTT 362
Qy 101 GlyMetAsnTyrrGlySerTyrrMetGluGlyLysHisIleProProProAsnMetThrThr 120
Db 363 GGATGACATATGAGAGCTTACATGAGAGAGAGATATTCGCGCTCCAAATATGACAACC 422
Qy 121 AsnGluArgValIleValProAlaAspProthrgLeuTrrPrrSerThrAspHisValArg 140
Db 423 AATGACGAGAAAGTATTGTGCGCAGCATCTTCAGTATGAGACGACAGCACTGTACGG 482
Qy 141 GlnTrrLeuGluTrrPAlaValLysGluTyrrGlyLeuProAspValAspIleLeuLeuPhe 160
Db 483 CAGTGGCTGGAGTGGCGAGTGAAGAGATGTTCTCCAGACGTGGACATCTTGTGTTTC 542
Qy 161 GlnAsnIleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 543 CAGAAACATTTGATGGGAAAGAGTGTGTAATAATGACCAAAAGATGACTTCCAGAGACTCAGC 602
Qy 181 ProSerTyrrAsnAlaAspIleLeuLeuSerHisLeuHisIleTyrrLeuArgGluArgLysAla 200
Db 603 CCGAGCTTAAACGAGATATCTCTCTGTCACACCTACCTACCTGAGAGAGAGAGAGACC 662
Qy 201 ThrPheIlePheProAsnThrSerValTyrrProGlnAlaThrGlnArgIleThrThrArg 220
Db 663 ACTTTTATTTTCCAAATACATCACTTTTACCCAGAGACAGCAAAAGATTAACAACAAG 722
Qy 221 ProAspLeuProTyrrGluGlnAlaArgSerAlaTyrrPrrSerHisSerHisProThr 240
Db 723 CCAGTATTACCTTATGACCAAGCGAGATCAGCGTGAAGAGACACAGCCATCCCACT 782
Qy 241 GlnSerLysAlaThrGlnProSerSerSerThrValPolyThrGluAspGlnArgPro 260
Db 783 CAGTCAAAAGCTACCCCAACATCACTTCAACAGGCCCAAAAGAGACCGAGCTCT 842
Qy 261 GlnLeuAspProTyrrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
Db 843 CAGTTAGATCTTATCAGATCTTGTGACCGACGACGACCGCTTGCATTCACAGGAGT 902
Qy 281 GlyIleIleGlnLeuTrrPglPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsn 300
Db 903 GGGCAGATACAGCTATGACAGTCTCTACTGAGGCTTGTGCGACAGCTCCAACTCCAAAC 962
Qy 301 CysIleThrTrrPglGlyThrAsnGlyLysLeuLysMetThrAspProAspGluValAla 320
Db 963 TGCATACCTGAGGAGGACAAATGGGAGTTCAAAGATGACAGACCTGATGAGATGGCT 1022
Qy 321 ArgArgTrrPglGlyArgLysSerLysProAsnMetAsnTyrrAspLysLeuSerArgAla 340
Db 1023 CGCGCTTGGGAGAGAGAGAAAGCAACCTTAACATGAACTTGAACAACTACCGCTGCA 1082
Qy 341 LeuArgTyrrTyrrAspLysAsnIleMetThrLysValHisGlyLysArgTyrrAlaTyrr 360
Db 1083 CTTCGCTACTACTATGACAAATAATATTATGACTAAAGTTCATGGTAAACGGCTATGCCCTAC 1142
Qy 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSerMet 380
Db 1143 AAAATTTGATTTCCAGGAATGGCTCAGGCGCTCCAGCTCCAGCTCCGAATTCATCCATG 1202
Qy 381 TyrrLysTyrrProSerAspLeuProTyrrMetSerSerTyrrHisAlaHisProGlnLysMet 400
Db 1203 TACAATATCCATGACAGACTCCCTACATGATGTTCTTCAATGACACACCCCAAGATG 1262
Qy 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420
Db 1263 AACCTTGTAGCTCCCAATCCCGCTTGGCTTGGCCGTAACCTGACAGCTTTTGTGGCTGCC 1322

QY 421 ProasnpProTyrTPasnsSerProthrglylyleTyrProasnthArgLeuProAla 440
 |||||||
 Db 1223 CCTAATCCATACCTGGAATTCACCAACTGAGGACATCCCAATACAGGCTCCAGACT 1382
 |||||||
 QY 441 AlaHisMetProSerHisLeuGlyTyrTyr 451
 |||||||
 Db 1383 GCTCATATGCTTCCATCTTGACCACTACTAC 1415
 |||||||

RESULT 3
 GSERG 1516 bp mRNA linear VRT 27-APR-1995
 LOCUS G.gallus ERG mRNA.
 DEFINITION X77159
 ACCESSION X77159
 VERSION X77159.1 GI:790439
 KEYWORDS ERG gene.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 1516)
 Dhordain, P., Dewitte, F., Desbiens, X., Stehelin, D. and
 Duterque-Cogulland, M.
 Mesodermal expression of the chicken erg gene associated with
 precartilaginous condensation and cartilage differentiation
 Mech. Dev. 50 (1), 17-28 (1995)
 95329425
 7605748
 2 (bases 1 to 1516)
 Duterque-Cogulland, M.
 Direct submission
 Submitted (17-JAN-1994) M. Duterque-Cogulland, CNRS UA 1160,
 Oncologie Molculaire, Institut Pasteur, 1 rue Calmette, 59019
 Lille, FRANCE

FEATURES
 source location/qualifiers
 1..1516
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /tissue_type="spleen"
 /clone_id="lambda gt10"
 /dev_stage="adult"
 63..1499
 /gene="ERG"
 63..1499
 /gene="ERG"
 /codon_start=1
 /protein_id="CAA54404.1"
 /db_xref="GI:790440"
 /translation="MASTIKELSVSESDSLFECAYGSPHLAKTEMTASSSEYGGT
 SKSPRPVQDDWLSQPARVYIKKCNPNQVNGSRNSPDDCSVAKGKMWSSSDVGM
 NYSYMEKHIIPPNNMTNRRVIVPADPILMSTDHYRQMLENAVKYGLPDVILF
 ONIDGKELCKMTKDDFORLTPSYNADILSLHLRYETPLPHLTSDDVDALQNSPL
 MHAIRNGATFIRPNTSVPEAQRITRRLDLYEQRARSAWTSHPROSKATOPS
 STYKTEQDRPQDLYOILISPTSSRLNRPSSGQIQLMQFLLEILSDSNSNCTMBET
 NGEFKMTDPEVARFRWGERKSKPMNMDKLSRALRYTDKNIMTKVGRKRYAKFEDH
 GIAQALQHPHESMKYPSDLPYMSYVNAHPQKMFNVAPHPALPVTSSSFPAHP
 YWNSPTGIGIPNRLPAHMPSHLGTYY"

BASE COUNT 458 a 392 c 327 g 339 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,73e-159 Length: 1516
 Score: 2414.50 Matches: 450
 Percent Similarity: 94.14% Conservative: 0
 Best Local Similarity: 94.14% Mismatches: 1
 Query Match: 98.79% Indels: 27
 DB: 5 Gaps: 1

US-09-902-772-2 (1-451) x GSERG (1-1516)

QY 1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 |||||||

Db 63 ATGCCAACACTATTAAAGAACATATATCAAGTGTGATGAGACAGACCTCTGTTGAC 122
 |||||||
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrcGlnMetThrAlaSerSerSerGlu 40
 |||||||
 Db 123 TGTCCTTACGGATCGCCCTTGGCAAGACAGAAATGACAGCTCTCTTCCAGTGAA 182
 |||||||
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
 |||||||
 Db 183 TATGGCAAAACATCAAGAGATGAGCCCGCGGTTCCCGACGAGACGCTGTTATCAGACCCC 242
 |||||||

QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 |||||||
 Db 243 CCGGCGAGAGTTACACTTAAGATGAGTGAACCAACAGGTAAATGGGTCAAGGAAT 302
 |||||||

QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 |||||||
 Db 303 TCACCTATGACTGCACCGTGGCAAAAGAGGAAATGGTTACCACTTCAGACAAATGTT 362
 |||||||

QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120
 |||||||
 Db 363 GGGATGAACATATGGAACTACATGAGAGAGAGCATATTCGCGCTCCAAATATGACAAACC 422
 |||||||

QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrPheSerThrAspHisValArg 140
 |||||||
 Db 423 AATGACGAAGAGTATTTGTGCCAGCATCTTACGTTATGAGACACAGACCATGTATCGG 482
 |||||||

QY 141 GlnTyrLeuGluTyrAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
 |||||||
 Db 483 CAGTGGCTGGAGTGGGAGTGAAGAGATATGCTTCCACACGTCGACATCTTGTCTTC 542
 |||||||

QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 |||||||
 Db 543 CAGAACATGATGAGGAAGAGTGTGTAAATGACCAAAAGATGACTCCAGAGACATCAGC 602
 |||||||

QY 181 ProSerTyrAsnAlaAspIleLeuSerHisLeuHisTyrLeuArgGluArg----- 198
 |||||||

Db 603 CCGAGCTATACGAGATATCTCTCTGTACACCTTACATCACTCCAGAGACTCTCTTT 662
 |||||||

QY 198 ----- 198
 |||||||

Db 663 CCACATTGACTCAGATGATGTTGATAGAGCCCTTACAAACACTCCAGCGTTAATGCAT 722
 |||||||

QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213
 |||||||

Db 723 GCTGAAGAACACAGAGAGGACGACCTTATTTTCCAAATACATCAGCTTACCCAGAGCA 782
 |||||||

QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnIleAlaArgAspSerAlaTyr 233
 |||||||

Db 783 ACGCAAGATATACACAGAGCCAGATTTACTTATGAGCAAGAGAGAGATCAGCGCTGG 842
 |||||||

QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253
 |||||||

Db 843 ACGAGTCACAGCCATCCCACTCACTCAAAAGATACCAACCATCATCTTCAACAGTGGCC 902
 |||||||

QY 254 LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
 |||||||

Db 903 AAAACAAACAAACAGCGCTCTCTCACTTATGATCATTAGATCTTGTGACCGACGAGAGC 962
 |||||||

QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeuLeu 293
 |||||||

Db 963 CGTCTTGCAATATCCAGGAGTGGCGAGATACACTATAGGACGATCTCTACGTGGAGCTTCTG 1022
 |||||||

QY 294 SerAspSerSerAsnSerAsnCysIleThrTyrGlnGlyThrAsnGlyGluPheLysMet 313
 |||||||

Db 1023 TCGGACAGCTCCAACTCACTCATCTGCTGGAGGCGCAAAATGGGAGTTCAAGATG 1082
 |||||||

QY 314 ThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsn 333
 |||||||

Db 1083 ACAGACCCCTGATGAAGTGGCTCGCGCTGGGAGAGAGCAAAAGCAACCTACATGAAC 1142
 |||||||

QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
 |||||||

Db 1143 TATGACAACTCAGCGCTGCTGCTTCTACTATGACAAAATATATATGACTAAAGTT 1202
 |||||||

QY 354 H1SGLYLSARGTYRATATYLYSPHEASPHENH1SGLYL1EALAGNALALEUINPro 373
 DB 1203 CATGTTAAACGCTATACCTACAAATTGATTTCCACGGAATGCTCAGGCTTCCAGCT 1262
 QY 374 H1SPROPGIUSerSerMetYrlySfYrProSerAspLeuProYrMetSerSerYr 393
 DB 1263 CACCTCCAGATCATCATGATGTAACAATACCATGACGCTCCCTACATGATGTTCCATAC 1322
 QY 394 H1SAHAH1SPROGILNLMetAsnHeValAlAPROH1SPROPRoAlaLeuProValThr 413
 DB 1333 CATGACACACCCCAAGATGAACTTTAGCTCCCAATCCCTTTGCGCTTAAC 1382
 QY 414 SerSerSerPhePheAlaAlAProAsnProYrTrpAsnSerProThrGlyGlyThr 433
 DB 1383 TCATCCAGCTTTTGGCTGCCCCATATCCATACCTGAAATTCACCACTGGAAGCATCTAC 1442
 QY 434 ProAsnThrArgLeuProAlaAlaH1MetProSerH1SLeuGlyThrYrYr 451
 DB 1443 CCCAATACCAAGCTGCCAGCTGCTCATATGCTTCCATCTTGACACTACTAC 1496
 RESULT 4
 E31254
 LOCUS E31254 1528 bp DNA linear PAT 18-JUN-2001
 DEFINITION protein having cell calcifying inhibitory activity and gene
 encoding the same.
 ACCESSION E31254
 VERSION E31254.1 GI:13025686
 KEYWORDS JP 199075871-A/2.
 SOURCE unidentified.
 ORGANISM unclassified.
 unclassified.
 1 (bases 1 to 1528)
 REFERENCE Hiroyasu, I., Yoshinobu, H., Marijo, P., Joel, R. and Helena, E.
 AUTHORS Protein having cell calcifying inhibitory activity and gene
 TITLE encoding the same
 JOURNAL Patent: JP 199075871-A 2 23-MAR-1999;
 CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF
 PENNSYLVANIA
 COMMENT OS Unidentified
 PN JP 199075871-A/2
 PD 23-MAR-1999
 PE 29-MAY-1998 JP 1998166076
 PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI
 HIROYASU IWAMOTO, YOSHINOBU HIGUCHI, MARIJO PASHIPIKI, PI JOEL
 ROZENBROOM,
 PI HELENA E
 PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00,
 PC C12P21/02,
 PC C12P21/08, C12N15/00, A61K37/02
 CC Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..1528
 FEATURES Location/Qualifiers
 source 1..1528
 1..1528
 /organism="Unidentified".
 /db_xref="taxon:32644"
 BASE COUNT 464 a 395 c 330 g 339 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.75e-159 Length: 1528
 Score: 2414.50 Matches: 450
 Percent Similarity: 94.14% Conservative: 0
 Best Local Similarity: 94.14% Mismatches: 1
 Query Match: 98.79% Indels: 27
 DB: 6 Gaps: 1
 US-09-902-772-2 (1-451) x E31254 (1-1528)
 QY 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20

DB 63 ATGCACACACTATTAAGGAAGCATTAACAGTGGAGGAAAGACCACTGTTGGAG 122
 QY 21 CysAlaTYrGlySerProH1SLeuAlaYThrGluMetThrAlaSerSerSerGlu 40
 DB 123 TGTGCTTGGATGATCCGCCACCCTTCCAAAGACGAAAGACAGCCTCTCTTCCAGTGA 182
 QY 41 TYrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerInPro 60
 DB 183 TATGGCAACATCAAGATGAGCCCGCGCTTCCCGACGAGCTGTTATCACAGCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 243 CCGGCACAGTTACCATTAAGATGAGGTAAACCAACCAAGTTAAAGGGTCAAGAAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 303 TCACCTGATGACTGACGCGTGGCAAAAGAGGAAATGTTAGCAGTTCAACAAATGTT 362
 QY 101 GlyMetAsnTYrGlySerTYrMetGlnGluLysH1S1LeuProProAsnMetThrThr 120
 DB 363 GGGATGAACATAAGAGCTTACATGGAAGAGAGCATATTCGCTCCAAATATGACAACC 422
 QY 121 AsnGluArgValIleValIleProAlaAspProThrLeuTrpSerThrAspH1SValArg 140
 DB 423 AATGACGAAGAGTTATTTGTCGACACATCTCTAGTTATGAGACACAGCATCTACGG 482
 QY 141 GlnTrpLeuGluTrpAlaValLysGluTYrGlyLeuProAspValAspIleLeuLeuPhe 160
 DB 483 CAGTGCCTGAGATGGGCGACGTAAAGATGATGCTTCCAGACGCTGACATCTTGTGTTCC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 543 CAGAACATTTGATGGAAAGAGTGTGTAAATGACCAAAAGATCTTCAGAGACTCAGC 602
 QY 181 ProSerTYrAsnAlaAspIleLeuLeuSerH1SLeuH1SYrLeuArgGluArg----- 198
 DB 603 CCGAGCTTAAACGACGATATCTCTCTGTACACACTACACTACAGAGAGACTCTCTT 662
 QY 198 ----- 198
 DB 663 CCACATTTGACTTCAGATGATGTTGATAGCGCTTACAAACCTCCACGGTTAATGCAT 722
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTYrProGlnAla 213
 DB 723 GCTAGAAACACAGAGGAGGACCACTTTATTTTCCAAATACATCAGTTTACCACAAACA 782
 QY 214 ThrGlnArgIleThrThrArgProAspLeuProTYrGlnGlnAlaArgSerAlaTrp 233
 DB 783 ACGCAAGAAATTAACACAAAGCCACAGATTACCTTATGACCAAGCAGAGATCAGCGTGG 842
 QY 234 ThrSerH1S1SerH1S1ProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253
 DB 843 ACGAGTCCACAGCCATCCACTCAGTCAAGCAAAACCTACCAACATCATCTTCAACAGTGGCC 902
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTYrGlnIleLeuGlyProThrSerSer 273
 DB 903 AAAACAGAGAACCAAGCGTCTCAGTTAGATCTTATGATCTTGGACCAACAGCAGCAGC 962
 QY 274 ArgLeuAlaAsnProGlyLysGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeu 293
 DB 963 CGTCTTGCAAAATCCAGGAGAGTGGCAGATACAGCTATGAGCACTTCTGAGACTTCTG 1022
 QY 294 SerAspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyLuphelysMet 313
 DB 1023 TCGGACAGCTCCAACTCCAACTGATCACCCTGGGAGGCGCAAAATGGGAGTTCAAGATG 1082
 QY 314 ThrAspProAspGluValAlaArgArgTrpGlyLysArgLysSerLysProAsnMetAsn 333
 DB 1083 ACGAGCCCTGATGAGAGTGGCTCGGCGTTGGGAGAGAGGAAACCAACCTTAACATGAAAC 1142
 QY 334 TYrAspLysLeuSerArgAlaLeuArgTYrTYrTYrAspLysAsnIleMetThrLysVal 353

Db 1143 TATGACAAACTGACCGCTGACTTGGCTACTACTATGACAAATAATTATGACTTAAGTT 1202

Qy 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373

Db 1203 CATGTAAACGCTTGGCTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCAGGCT 1262

Qy 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393

Db 1263 CACCTCCAGAAATATCATGTACAAATATACCATCAGACCTCCCTCATGATGATTCCTAC 1322

Qy 394 HisAlaHisProGluLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413

Db 1323 CATCACACCCCGCAGAAAGATGAACTTGTACTCCCATCCCTGCTTTCGCCGTAAAC 1382

Qy 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyr 433

Db 1383 TCATCCAGCTTTTGGCTGGCCCTAATCCATACATGAAATTCACCAACGTGAGGATCTAC 1442

Qy 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451

Db 1443 CCCAATACCAAGCTGCTCAGCTGCATATGCTTCCATCTTGGCACCTACTAC 1496

RESULT 5

AB073080 2209 bp mRNA linear ROD 17-OCT-2001

LOCUS AB073080

DEFINITION Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,

ACCESSION AB073080

VERSION AB073080.1 GI:16197544

KEYWORDS

SOURCE Mus musculus

ORGANISM Mus musculus cDNA to mRNA, clone:Erg-3.

REFERENCE 1

AUTHORS Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.

TITLE Mus musculus Erg mRNA

JOURNAL Published Only in Database (2001)

REFERENCE 2 (bases 1 to 2209)

AUTHORS Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gscc.riken.go.jp, URL: http://hgp.gscc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

FEATURES

source location/Qualifiers

1..2209

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="16"

/clone="Erg-3"

1..2209

/gene="Erg"

148..1608

/gene="Erg"

/note="homolog of human ets-related gene (ERG), transcript variant 3"

/codon_start=1

/protein_id="BAB69950.1"

/db_xref="GI:16197545"

/translation="MIQTWLPDPAARIHKEALSVSEDSLEFCAATGPHLAKTEMTASS

SSDYGSOTSKSPRYPOODWLSOAPARVITIMECPNSQVNSRNSPDCSVNKGKMG

SPDTVSGSYGMEKEKHPNMTNRRVIVPADPTLMSDHYRQLEMAVEXGL

DVDVILFQNDKEELCKMTKDDFORLTPSRNADLLSHLYLRTPLPHLTSDDVDA

LQNSPRLHARNTGGAAIFPNTSVPEATQRTTRDLYEPERRASAWGSHSLYQA

SKAASPSPAVPTEDRQDLPQIIGPTSSRLANGSGQIQIMQPLLELSSNS

NCITGCTNEFKMTDDEVARMGERSKPNMYDKLSALRLYYDKNLTLYGSR

YAFEDFHGIAOLAPHPPESSLYKPYSDPYMGSGYAHQKNMFVPHPALPVTSS

SFSPSPRYNNSPGGITYPNTRLPASHMPSHLGTY"

BASE COUNT 606 a 608 c 545 g 450 t

ORIGIN

Alignment Scores:

Pred. No.: 5,18e-149 Length: 2209

Score: 2267.00 Matches: 421

Percent Similarity: 90.59% Conservative: 12

Best Local Similarity: 88.08% Mismatches: 17

Query Match: 92.76% Indels: 28

DB: 10 Gaps: 2

US-09-902-772-2 (1-451) x AB073080 (1-2209)

Qy 2 AlasSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21

Db 172 GCACCTCATATTAAAGAGGCGCTTGTCACTGTGTGACGAGCACCGTACTATTATGAGTGT 231

Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41

Db 232 GCCTACGGAAAGCCACACCTGGCTAAGACACAGATGACCCGATCTTCCAGTGACTAT 291

Qy 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnProPro 61

Db 292 GGCACAGCATCCAGATGAGAGTCCAGAGTCCCTCAGCAGAGACTGGCTGTCTCAAGCCCA 351

Qy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81

Db 352 GCCAGGCTCACCATCAAGATGAGATGCAACCCCTGATGAGTGGTTCAGACTCA 411

Qy 82 ProAspAspCysSerValAlaLysGlyGlyMetValSerSerSerAspAsnValGly 101

Db 412 CCTGATGATGTCACATGTGAACAAAGGTGGAGATGGTGGACCCCGGATCTGTGGG 471

Qy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121

Db 472 ATGAGTACGCGCACCTCATGTGAGAGAGATGATGCGCGCTCCCAATATGACCAAAAT 531

Qy 122 GluArgArgValIleValProAlaAspProThrIleutPserThrAspHisValArgGln 141

Db 532 GAGCGCAGAGTGAATGCGTCCGAGATCTCACTGTGGAGACACAGCATGTCCGACAG 591

Qy 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161

Db 592 TGGCTGAGTGGGCGGCGGAAAGATATGGCTCTCATATGAGCGTCTTACTATTTCAG 651

Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181

Db 652 AATATGATGGGAAAGAGAGCTGTGCAAGATGCAAAAGATGACTTCCAGCGGCTCAGCCG 711

Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg----- 198

Db 712 AGCTACAAATGCCGACATTTCTCTCTCATCTCCATCTCAGACAGAGACTCCCTTCCA 771

Qy 198 ----- 198

Db 772 CATCTGACTTCGATGACGTTGATTAAGGCTTTACAAACTCTCCACGTTAATGATGCC 831

Qy 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 214

Db 832 AGAAACACAGGGGTGACGCTTATTTCCAAATCTTCAGATATCCGAAAGCTACG 891

Qy 215 GluArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThr 234

Db 892 CAAGAATTTACACTAGCCAGATTTACCTTATAGAGCTCCAGGAATATGAGCTTGAGC 951

Qy 235 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 253

Db 952 GGCACAGCCACCTCACCTCTCACTCCAAAGCTGCTAGCCATCTCCCTTGACAGTCCC 1011

Qy 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273

Db 1012 AAAAAGTAAAGCAGGCTCTCTGATGATCTTACAGATCTCGGAGACGACAGCAGTACG 1071

Qy 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeu 293

```

|||||
Db 1072 CGCCTTCTATCCAGGATAGGCGCAGATCCAGCTGTGGAGTCTCTGTCGAACTCCTG 1131
Qy 294 SerSpSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 313
Db 1132 TCACAGCAGCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCT 1191
Qy 314 ThrAspProAspGluValAlaArgArgTropGlyGluArgGlySerSerSerSerSerSer 333
Db 1192 ACACACCCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1251
Qy 334 TyrAspLysLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 353
Db 1252 TATGACAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1311
Qy 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyLeuAlaGlnAlaLeuGlnPro 373
Db 1312 CACGGGAGCGCTACGCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCT 1371
Qy 374 HisProProGluSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 393
Db 1372 CACCTCTCTAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAT 1431
Qy 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
Db 1432 CACGCCACCCCGAGAGATGAACTTGTGTCTCCACCTCCACCTCCACCTCCACCTCCAC 1491
Qy 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrPheSerSerProThrGlyGlyIleTyr 433
Db 1492 TCTTCACGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1551
Qy 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyTyrTyrTyr 451
Db 1552 CCGAACACTAGGCTCCACGACGACATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1605

RESULT 6
XLAJ4126 2012 bp mRNA linear VRT 21-JAN-2000
LOCUS Xenopus laevis erg gene (erg_E).
DEFINITION AJ224126 GI:5420047
ACCESSION AJ224126.1 GI:5420047
VERSION 1
KEYWORDS Erg gene; transcription factor.
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 2012)
AUTHORS Baltzinger M., Mager-Heckel A.M. and Remy P.
TITLE Xl erg: expression pattern and overexpression during development
plead for a role in endothelial cell differentiation
Dev. Dyn. 216 (4-5), 420-433 (1999)
JOURNAL 20099678
MEDLINE 10633861
PUBMED 2 (bases 1 to 2012)
REFERENCE Baltzinger M.
AUTHORS Baltzinger M.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1998) Baltzinger M., UPR9005, MMDD, Centre
National de la Recherche Scientifique, 15, rue Rene Descartes,
67084, FRANCE
FEATURES
source
1..2012
/organism="Xenopus laevis"
/db_xref="taxon:8335"
/clone="lambda ZAPII-erg E"
/clone_id="lambda ZAPII"
/dev_stage="stage 30 embryo"
225..1595
/gene="erg"
225..1595
/gene="erg"
225..1595
/gene="erg"
/function="Transcription factor"
/codon_start=1

```

```

/protein_id="CA846567.1"
/db_xref="GI:5420048"
/db_xref="SPTREMBL:O9M629"
/translation="MIDTVDPDSHIREALSVSEDSLEFECTGPHLTKTENTASS
SSDYGQTSKMSPRVQDQWLSQPSRTITMECPNSOVNSRSPDSCSGKSKMG
GSDNVMNYSYIEEKHISPMNTNRRYIVPADPLMSTDIHRQWLEWAKYEGLP
DVAFLFONIDGKELCKMTKEDEFRLEPSYNAIDLSHLHYLRGASLFEPSVVO
DANRIRPSRODLSYEPSRSAMNHPPAPPSKASQSPSTTVKTEDEPRQDLPYOLIGPT
SSRLANGSQIOLOMFLLELSDSSSNKLTMEGTGERKMDPDEARWRKRSK
PNNMIDKLSRLARITIDKNIMTVHGRVATKEDFHGIAQALDHPPESTMTKPSLEL
PYMSYHAHPQKNMFVAHPHPLPVTSSEFPAPNAYWNSPTGSIYPNTRLPSHMS
HLGTYV"
BASE COUNT 634 a 479 c 395 g 504 t
ORIGIN
Alignment Scores:
Pred. No.: 7,84e-144 Length: 2012
Score: 2192.00 Matches: 402
Percent Similarity: 94.63% Conservative: 21
Best Local Similarity: 89.93% Mismatches: 22
Query Match: 89.69% Indels: 2
DB: 5 Gaps: 2
US-09-902-772-2 (1-451) x XLAJ4126 (1-2012)
Qy 5 IleYsgLualLeuSerValSerGluAspGlnSerLeuPheGluCysAlaTyrGly 24
Db 258 ATCAAGAGAGCTCTTTCAGTGAGTGAGTGAAGATAGCTGTTATTCAGAGTACCTACGGA 317
Qy 25 SerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyTyrGlyGlnThr 44
Db 318 ACGGCACATCTTACTAAAGGAGATGACCGCATCTTCCTCCATGACTATGGCAAAACC 377
Qy 45 SerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnProProAlaArgVal 64
Db 378 TCGAAATGAGTCCACCTGTGCTCCACCAAGAGACTGCTCCACAGCTCCATCCAGGGTG 437
Qy 65 ThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspAsp 84
Db 438 ACCATCAAGATGGATCAACCTTAGTCAAGTCAATGATCAAGAGACTCCCGGATGAC 497
Qy 85 CysSerValAlaLysGlyGlyLysMetValSerSerSerSerAspAsnValGlyMetAsnTyr 104
Db 498 TGCAGTATAGGAAAGAGACCAAAATGGGAGGCGTCGATATGTGGATGAATATTAT 557
Qy 105 GlySerTyrMetGluGlyLysHisIleProProProAsnMetThrThrAsnGluArg 124
Db 558 GCGAGCTACATAGCAAGAAAACACATTCTCCCAAAACATGACCAACCAATGAGAGAGA 617
Qy 125 ValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgIleTyrLeuGlu 144
Db 618 GTGATTTCCTGCTGACCTACTCTTGTGAGCACAATATGATGACACAGTGGCTGAG 677
Qy 145 TrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAsp 164
Db 678 TGGCAGATATAAAGAAATACGGCTTCCAGATGTGATGCTGCTGTTCCAAACATGAC 737
Qy 165 GlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsn 184
Db 738 GGCAGGAATTATGACAGATGACCAAGAAAGATTCCATGACACACGACCAAGCTACAAAT 797
Qy 185 AlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThrPheIlePhe 204
Db 798 GCTGATATCTTCTCTGCAATCTTCACTACCTCAGAGAAAGAGTGGCTCTTATATTTC 857
Qy 205 ProAsnThrSerValTyrProGluAlaThrGlnArgGlyLeuThrArgProAspLeuPro 224
Db 858 CCAAACTCATCCGTTATCAAGACGCAAAAGAAATCCCAAGACGCAAGATTATCA 917
Qy 225 TyrGluGlnAlaArgArgSerAlaTyrThrSerHisSerHisProThrGlnSerLysAla 244
Db 918 TATGAGCCCTCAAGAGATCTGCTGAGCAAAATCATCTCAGCA--CCATCAAAAGCA 974

```


QY	245	ThrglnProSerSerSerThValProLysThrGluuAspGlnArgProGlnLeuAspPro	264
Db	975	TCACACCA---TCTACCAAGTTCCAAAACAGAACCCAGGCGCACGCTAATCTT	1031
QY	265	TyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGln	284
Db	1032	TATCAGATTCTTGGGCGCAACAGACGCCGACTTGGAATTCAGAGAAAGTGGCAGATTGAG	1091
QY	285	LeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrp	304
Db	1092	CTCTGGCAATCTTACTGGAACGTGTTTGGATAGTGGTCCAACTCCAACTGCATTACTTGG	1151
QY	305	GluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAlaArgArgProGly	324
Db	1152	GAAGGACCAATTGAGGAATTTAAGATGACCCATCTTGATGAAGTGGCAAGACGTTGGGGG	1211
QY	325	GluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyr	344
Db	1212	GAGAGGAACAACCAACCAACATGACATATATAAGCTCACCCGTCGACTTCGTACTAC	1271
QY	345	TyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPheAspPhe	364
Db	1272	TACGATAAAATATATATAGTAAAGTCCATGGCAAGCCGTATGCTTACAAATTTGATTTT	1331
QY	365	HisGlyIleAlaGlnAlaLeuGlnProHisProProGluSerSerMetTyrLysTyrPro	384
Db	1332	CATGGAGATGGTCACAGCTCTTCAACCTCATCTCTCCAGATACATCATATATACAGATCCC	1391
QY	385	SerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMetAsnPheValAla	404
Db	1392	TCAGAACTTCATATATACGCTCATACATGCATCATCCACAGAAAGATAACTTTGTTGCC	1451
QY	405	ProHisProProAlaLeuProValThrSerSerSerPhePheAlaAlaProAsnProTyr	424
Db	1452	CCCCATCCGCGCAGCTTTACTCGTGACATCGTCAAGTTCTTTGGCAGCCCTAATGCAATAC	1511
QY	425	TrpAsnSerProThrGlnGlyTyrIleTyrProAsnThrArgLeuProAlaAlaHisMetPro	444
Db	1512	TGGAAATTCACCAACTGGGAATTTATTCCAATACTCGGCTGGCCAGCTAAGCCATATGCT	1571
QY	445	SerHisLeuGlyThrTyrTyr 451	
Db	1572	TCTCATCTTGGACCTACTAT 1592	
RESULT 7			
HUMERG2			
LOCUS	HUMERG2	3166 bp ss-RNA	linear PRI 08-NOV-1994
DEFINITION	Human erg2 gene encoding erg2 protein, complete cds.		
ACCESSION	M17254		
VERSION	M17254.1 GI:182186		
KEYWORDS	erg 2 protein.		
SOURCE	Human cell line COLO 320, cDNA to mRNA, clone lambda 12.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	1 (bases 1 to 3166)		
TITLE	Rao,V.N., Papas,T.S., and Reddy,E.S.		
JOURNAL	erg, a human ets-related gene on chromosome 21: alternative		
MEDLINE	splicing, polyadenylation, and translation		
PUBMED	Science 237 (4815), 635-639 (1987)		
FEATURES			
source	3299708		
	Location/Qualifiers		
	1..3166		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="21q22.3"		
	257..1645		
	/gene="ERG"		
	257..1645		
	/gene="ERG"		
	/note="erg 2 protein"		
	/codon_start=1		

		repeat_region	291..298	/note="8-bp repeat sequence"	
		repeat_region	509..516	/note="8-bp repeat sequence"	
		polyA_signal	3143..3149	/note="putative"	
	BASE COUNT	915 a	770 c	768 g	713 t
	ORIGIN	227 bp upstream of PvuII site; chromosome 21.			
	Alignment Scores:				
	Pred. No.:	6,69e-143	Length:	3166	
	Score:	2182.00	Matches:	407	
	Percent Similarity:	91.72%	Conservative:	14	
	Best Local Similarity:	88.67%	Mismatches:	24	
	Query Match:	89.28%	Indels:	14	
DB:		9	Gaps:	4	
US-09-902-772-2 (1-451) x HOMER2 (1-3166)					
QY	2	AlaserthrlllelysglialaleuservAlvalSerlglnaspginSerleupheglucys	21		
DB	281	GCAGCTCATTCACAGAGAGACCTTATCATGTTGTAGTGCAGACCGATGCTTGTAGAGT	340		
QY	22	AlatryglSerprroHlileuAlalyshtrolumethrAlaserSerSerleutyr	41		
DB	341	GCTTACGGAGACGCCACACCTGGCTTAAGACAGATGACCGGCTCCTCCAGGACTAT	400		
QY	42	GLyglntHrserlysmetSerProarvalProglnglnaspTyrpleSerGlnProPro	61		
DB	401	GGACAGACTTCCACAGATGAGCCACCGCTCCCTCAGCAGATTTGGCTGTCTCAACCCCCA	460		
QY	62	AlaarvalthrlllelysmetglucysanProasnlglnAlaanglySerargasnSer	81		
DB	461	GCCAGGGTCCACCATCAATAATGGAATGTATACCTTGCCAGGTGATGGCTCAAGCAACTCT	520		
QY	82	ProaspaprcysservAlAlalySGlyglylysmetvalSerSerSerAspAsnvalgly	101		
DB	521	CCTGATGATCAGTGTGGCCAAAGCGCGGAAGTGTGGGACACCCACGACACCGTTGGG	580		
QY	102	MetAsnTyrGlySerTyrMetGlnGluysHlsileProProProasnMetThrThrAsn	121		
DB	561	ATGAACTACGCCAGCTACATGAGAGAGACATGCCACCCCAACATGATACCAACGAC	640		
QY	122	GLuarGarvalAllevalProAlaAspProthrleutTyrSerThrAspHisvalargln	141		
DB	641	GAGGCGACAGTTATTCGTCGCAGAGATCCTACGGTATGAGACTACAGACCATGTCGGGAG	700		
QY	142	TrpleuglutrPalavallysglutrtyglyleuProaspvalAspilleuleuPhegln	161		
DB	701	TGGCTGGAGTGGGGGCGTGAAGAATATGCGCTTCCACACGCAACATCTGTATTCAG	760		
QY	162	AsnllleasprglylsgluleucyaleuymethrThrlyrsasprheglnargleutThrPro	181		
DB	761	AACATTCAGTGGGAAGAACTGTGCAGATGACCAAGACGACATTCACAGAGCTCACCCCC	820		
QY	182	SerTyrAsnAlaAspilleuleuSerHisleuHslyrleuArgGluargGlyAlaThr	201		
DB	821	AGCTACAAAGCGGCACATCCTTCTCTCATCTTCACATCTCAGACGAGACTCCCTTCCA	880		
QY	202	PheillepheProasnThrSerValTyrProGluAlaThrGlnargGlyleuThrArgPro	221		

Db	881	CATTGG-----ACTTCA-----GATGATGTTGATTAAGACCTTACAAACACCTOCA	925
OY	222	-----AspleupProTyrgluGlnAlaIArgSerAlaTrp	233
Db	936	CGGTTAATGCATGCTAGAAACACAAATTTACCATATGAGCCCCCGAGAGATCACCTGG	985
OY	234	ThSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal	252
Db	986	ACCGGTACAGGCGCACCCACCGCCCAAGTCGAAAGCTGCTCAACCATCTCTCCACAGT	104
OY	253	ProLysThrGluAspGlnArgProGlnLeuAspProTyrglnIleLeuGlyProThrSer	272
Db	1046	CCCAAACTGAAAGACACAGCTCTCAAGTTAGATCTCTTACGATTCCTTGAGCAACAAGT	1105
OY	273	SerAlaGluAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeu	292
Db	1106	AGCCGCTTGCAAAATCCAGGACGTGGCCAGATCCAGCTTTGGCAATTCCTCTGGAAGTC	1165
OY	293	LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLeuPheLys	312
Db	1166	CTGTGCGACAGCTCCAACTCCAGCTGCATACCTGAGGAAGCACCACGGGGAGTTCAAG	1222
OY	313	MetThrAspProAspGluValAlaIArgArgTrpGlyGluIArgLysSerLysProAsnMet	332
Db	1226	ATGACGAGATCCCGACGAGGTGGCCCCGGCGCTGGGGAGAGCGGAAGCAAAACCAACTG	1285
OY	333	AsnTyraAspLysLeuSerAlaLeuArgTrpTyrrTyraAspLysAsnIleMetThrLys	352
Db	1286	AACATGCATTAAGCTCAGCGCGCGCCCTCGGTACTACTCTATGCACAAACATCATATGACCA	1344
OY	353	ValHisGlyLysArgTrpTyraIleTyrrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln	372
Db	1346	GTCCATGGGGAAGCGCTACGCCCTACAAAGTTCGACTTCACAGGGAGTCCGACAGCCCTCCAG	1405
OY	373	ProHisProProGluSerSerMetTyrrTyrrProSerAspLeuProTyrrMetSerSer	392
Db	1406	CCCCACCCCCCGGAGTCACTCTGTGCACAGTACCCCTAGACCTCCGCTACATGGGGCTCC	1465
OY	393	TyrrHisAlaHisProGlnLysMetAsnPheValAlaIProHisProProAlaLeuProVal	412
Db	1466	TATCAGCGCCACCCACACAGATGAACTTTGTGGCGCCCAACCTCCAGCCCTCCCGGTG	1522
OY	413	ThrSerSerSerPhePheAlaIAlaIProAsnProTyrrTrpAsnSerProThyGlyLysIle	432
Db	1526	ACATCTTCAAGTTTTTTGTGCTGCCCAAAACCCATCACTGGAATTCACCAACTGGGGGTATA	1585
OY	433	TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrrTyrr	451
Db	1586	TACCCCACTAGAGCTCCCAACAGCAGCATATGCTCTTCATCTGGGACATTACTAC	1642
RESULT 8			
AB031088			
LOCUS	AB031088	1808 bp	mRNA
DEFINITION	Rattus norvegicus VESPI4 mRNA for vascular endothelial cell		
ACCESSION	AB031088		
VERSION	AB031088		
KEYWORDS	specific protein 14, complete cds.		
SOURCE	AB031088.1 GI:15128488		
ORGANISM	Rattus norvegicus		
	vascular endothelial cell specific protein 14.		
	Rattus norvegicus liver cDNA to mRNA.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 1808)		
AUTHORS	Aoki,T., Toyoda,H., Nishimoto,S., Tawara,J., Ukai,Y. and		
	Kumurasaki,T.		
TITLE	Identification of VESPI4,a vascular endothelial cell specific		
	protein		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1808)		
AUTHORS	Aoki,T., Toyoda,H., Nishimoto,S., Tawara,J., Ukai,Y. and		
	Kumurasaki,T.		
TITLE	Direct Submission		

Submitted (12-AUG-1999) Tomomi Aoki, Taisho Pharmaceutical Co., Ltd., Medicinal Research Laboratories, Molecular Biology Lab.; 1-403, Yoshino-cho, Omiya, Saitama 3308530, Japan (E-mail: ai1704@com.taisho.co.jp, Tel: 048-663-1111, Fax: 048-652-7254)

FEATURES
source Location/Qualifiers
 1..1808
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /cvsue_type="Liver"
 1..1808
 /gene="VESE14"
 63..1430
 /gene="VESE14"
 /note="similar to human egr1"
 /codon_start=1
 /product="Vascular endothelial cell specific protein 14"
 /protein_id="BAB62744.1"
 /db_xref="GI:15128489"
 /translation="MASTIKELALSVEKSDSLFECAATGPHLAKEWTMASSSDPYGOTSKMSPYODDMLSDPPARYTIKMECSNPOVNGSRNSDSCSYKGGKAMGSDPTVGM SYGSMEEKHMPRPNMNTNERARYVAPDPLMTSDYHROMLENAVKRYGLLDVYLE ONIDGELCKTKDFORLTPSYNADLLSLHLRLRETP.LPHLPSDDVDAIANSPLL MHAANTDLPYEPPIRSTWTGSHSPTPOSKAOPSPTVPTEDORQLDYLQILGPS SRLNPSGGQIQIOWDFLELLSDSNSNCITWGTGNEFMTDPEDEARWGRKSKP SMTANDKSRLAIRYYDQFLEINIMTKVGRKAYAKFPFHGIAOLAPHPSSLYKPSDLP YMGSYTHNPQKMNFAVRPALPVTSSSFATRPVNPSTGGIYVPTRLPASHMPSH LGTYR"

BASE COUNT 510 a 522 c 426 g 350 t

ORIGIN

Alignment Scores:
 Pred. No.: 5.56e-143 Length: 1808
 Score: 2179.00 Matches: 406
 Percent Similarity: 91.30% Conservative: 14
 Best Local Similarity: 88.26% Mismatches: 26
 Query Match: 89.16% Indels: 14
 Gaps: 4

US-09-902-772-2 (1-451) x AB031088 (1-1808)

OY 1 MetLaseRhrlllelysglualaleuSerValValSerGlunaspInSerleuepHeGu 20
 |||||
 Db 63 ATGGCCAGCACTATTAAAGAGGCGCTTCTGAGTGTGAGCAAGGACCGATCAATTGGAG 122
 21 CysAlatYrGlySerProHsleuAlaLysThrGluMetHraLaseSerSerSerGlu 40
 |||||
 Db 123 TGTGCGTAAAGGAAAGCCCAACCTGTGGTAAAGCAAAATGACCGCATCTTCCAGTGAC 162
 OY 41 TyrGlyInThrSerLysMetSerProArGValProGlnGlnsPrrleuSerGlnPro 60
 |||||
 Db 183 TATGGCCAGCAATCCAAAGATGAGTCCAGAGTCCCAACAGCAGCTGGCTGTCAACCC 242
 61 ProAlaArGValThrIleLysMetGlnCysAsnProAsnGlnValAsnGlySerArGAsn 80
 |||||
 Db 243 CCAGCCAGGGTCCACCAATGAGATGAGATGCAACCCCTAGCCAGGTGAACGGTTCCAGGAAC 302
 OY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 |||||
 Db 303 TCACCTGACGCAATGACATGTGACCAAAAGGTGGGAAGATGTGGGCAACCCCTGATACTGTG 362
 OY 101 GlyMetAsnTrpGlySerTrpMetGlnGluLysHisIleProProProAsnMetThrThr 120
 |||||
 Db 363 GGAATGTAGCTTACGGCAGCTACATGAGAGGAAGACACATGCGGCCCCCAATTATGACCACG 422
 OY 121 AsnGluArGArGValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
 |||||
 Db 423 AATGACGCGACAGTGAATCGTCCCTGGCAGATCTTACTGTGTGAGACACAGCAATGTCCGG 482
 OY 141 GlnTrpLeuGlnTrpAlaValAlaLysGluTrpGlyLeuProAspValAspIleLeuLeuPhe 160
 |||||
 Db 483 CAGGGCTGAGAGTGGGACATGAAAGATTAAGGCGCTCTTACAGCTGACGCTTATTAATTT 542

QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 543 CAGAAATATGACGGGAGAGAGCTGTGCAGATGACAAAGATGACTTCCAGAGGCTACT 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGlyAla 200
 Db 603 CCGAGCTACAAAGCCGACATCTTCTCTCAGATCTCCAGTACCTCCAGAGAGACTCCCTT 662
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGlnIleAlaThrGlnArgIleThrThrArg 220
 Db 663 CCACATCTG-----ACATCC-----GACGACCTGTATAGAGCTTACAAAACTCT 707
 QY 221 Pro-----AspleuProTyrGluGlnAlaAlaArgSerAla 232
 Db 708 CCAAGGTAAATGCTAGCTAGAAACACAGATTTACTTACCTTACCTCCAGAGAGATCAACC 767
 QY 233 TrpThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThr 251
 Db 768 TGGACGGCCACAGCCACCCACCCCTCAGTCCAAACCTCCAGCATCTCCCTCCACA 827
 QY 252 ValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThr 271
 Db 828 GTGCCCAAAACTGAAGACGAGCTCCCATGATGATCTTACAGATCTCGGAGCCACC 887
 QY 272 SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuGlu 291
 Db 888 AGTAGTCGCTTCTATCCAGGTAGTGCCAGATCCAGCTGTGGCAGTTCTGCTAGAA 947
 QY 292 LeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyAlaPhe 311
 Db 948 CTCTGCTGTACAGCTCCAACTCCAACTGCATACCTGGGAGGACCAACGGGAGTTG 1007
 QY 312 LysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsn 331
 Db 1008 AAGATACAGAACCCGGGTAGTGCTGCGCGCTGGGGGAGGAGCAACCAACCCAAC 1067
 QY 332 MetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThr 351
 Db 1068 ATGACCTATGACAAACCTCAGCCGCTGCTGCTACTACACAAACAAATCATGACC 1127
 QY 352 LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeu 371
 Db 1128 AAGGTGACGGGAGAGGCTATGCTCAAGATTGACTTCACGGGATGCGCCAGGCCCTG 1187
 QY 372 GlnProHisProProGluSerSerMetTyrLysTyrProSerAspleuProTyrMetSer 391
 Db 1188 CAGCCCATCCGCCCGAGTGTCTCTGACAACTACCCCTCGACCTGCGCTACATGGGC 1247
 QY 392 SerTyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411
 Db 1248 TCCTATACACCCACCCCGAAGATGACTTGTGGCTCCACCCCTCGGCCCTCCCA 1307
 QY 412 ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGly 431
 Db 1308 GTCACATCTTCCAGTTCTTGTGTACCCGGAACCTACTGGAATTCGCCGATGGGGGC 1367
 QY 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
 Db 1368 ATCTACCGAACAATGAGTCCAGCCAGCCACATATCCCTTCACCTGGGACACTTATAC 1427
 RESULT 9
 AB073079 2133 bp mRNA linear ROD 17-OCT-2001
 LOCUS
 DEFINITION Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,
 transcript variant 2, complete cds.
 AB073079
 AB073079.1 GI:16197542
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE Mus musculus
 ORGANISM Mus musculus
 REFERENCE 1 Mus musculus CDNA to mRNA, clone:Erg-2.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.
 TITLE Mus musculus Erg mRNA
 JOURNAL Published Only in Database (2001)
 REFERENCE 2 (bases 1 to 2133)
 AUTHORS Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 FEATURES
 source location/Qualifiers
 1..2133
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"
 /clone="Erg-2"
 1..2133
 /gene="Erg"
 149..1540
 /gene="Erg"
 /note="homolog of human ets-related gene (ERG), transcript variant 2"
 /codon_start=1
 /protein_id="BAB69949.1"
 /db_xref="GI:16197543"
 /translation="MIQYVDPAAHIKEALSVSPOSLEFCAVGTPHLAKTEMTASS
 SSDYGTSKSPVPOODMISQAPARTTMECNPSQVNSRSPSPSCSYNKKAGV
 SPDTVMSTGSIIEKHVPPNPRTNERIVTADPTLSTDIHVRLVAVAYEGL
 DDDVLLFONIDGELCKMTKDDQRLTPSYNADILSHLYRETLPHLTSDVKA
 LONSPRIHARNRTGGAFTFPTNSVPEARTYTPPAOPSPSAVKEDQRPQDP
 YOLIGPTSSRLANRGGGLOIMQFLLELSDSSNSNCITMEGTGSEFKMDPDEVAR
 WGRKSKPMNNYDKLSRALRYDYDKNIMTKVHGKRYARFDFGIALOAPHPRESSL
 YKPSDLPYMGSYHAHPQKNFSPHPALPYTSSSFSPSPNPNYMSPTGIGTPNTRL
 PASHMSHLCTTY"
 BASE COUNT 587 a 579 c 531 g 436 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,55e-141 Length: 2133
 Score: 2159.50 Matches: 403
 Percent Similarity: 87.00% Conservative: 12
 Best Local Similarity: 84.49% Mismatches: 13
 Query Match: 88.36% Indels: 49
 DB: 10 Gaps: 2
 US-09-902-772-2 (1-451) x AB073079 (1-2133)
 QY 2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
 Db 173 GCAGCTCATATTAAGAGAGCGCTTGTCAAGTTGTAGCGAGGACGACACTATTGTAGCTG 232
 QY 22 AlaTyrGlySerProHisIleuAlaLysThrGluMetThrAlaSerSerSerGlyTyr 41
 Db 233 GCTTACGGAAGCGCAACCTGCTAGACAGAAATACCGACCTCTTCCAGTGAATA 292
 QY 42 GLYGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnProPro 61
 Db 293 GGCACACATCCAGATGAGTCCACAGAGTCCCTCCACAGAGATGCTGTCTCAAGCCCA 352
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 Db 353 GCAGAGGTACCATCAAGATGAGTCCACAGAGTCCCTCCACAGAGATGCTGTCTCAAGCCCA 412
 QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAsnValGly 101
 Db 413 CCTGATGAGTGCAGTGTGAACAAAGGTGGAAGATGTGGAGCCCGGATCTGTGGGG 472
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
 Db 473 ATGAGCTACGGACCTACTGAGGAGGAGCAAGCATGTGCGCTCCCAATATGACCAAAAT 532

Db	486	ATGAGTCCAGCGTGTGGCTCACAAGACACTGGCTCTCCAGGCTCCATCCAGGGTGAACATC	545
OY	67	LysMetGIuCyAsnProAsnGlnValAsnGlySerAlaGlnSerProAspAspCysSer	86
Db	546	AAGATGTGAATCAACCCCTAGTCAAGTCAATGGATCAAGAGACCTCCCGCATGACATGCAGT	605
OY	87	ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySer	106
Db	606	ATAGAAAAGGAGACCAAAATGGGTGGAGGTCGGATATATTGGCAATGAATTAATGCGACG	665
OY	107	TyrMetGIuGIuLysHisIleProProAsnMetThrThrAsnGluArgValIle	126
Db	666	TACATAGAAAGAAAAACACATTTCTCCCCCAACATGACAAACCAATGAGAGAGATGATT	725
OY	127	ValProAlaAspProThrLeuThrSerThrAspHisValArgGlnTrpLeuGlnTrpAla	146
Db	726	GTACCTGTGACCCCTACTTTGTGGAGCACATCATGTGCACAGAGTGGCTAGAGTGGGCA	785
OY	147	ValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys	166
Db	786	ATTAAGAAGATAGGGCTTCAGAGATGTGATGTGCTCCGTGCCAACAATCGAGGAGCA	845
OY	167	GluLeuCysLysMetThrLysAspAspPheGlnArgIleThrProSerTyrAsnAlaAsp	186
Db	846	GAATTATGCCAAGATGACCAAGAAAGATTTCATGCATGCACGACCAAGCTACAAATGCTGAT	905
OY	187	IleLeuLeuSerHisIleHisIleTyrLeuArgLysIle	198
Db	906	ATCCTTCTCTGCAATCTTCACTACTGTCAGAAACTCTCTCCACATTTGACTTCAGAT	965
OY	199	-----Gly 199	
Db	966	GATGTGATGAAGGCTTACAAAACCTCCACAGGTTAATGATGCTAGAAATATACAGAGAT	1025
OY	200	AlaThrPheIlePheProAsnThrSerValIleTyrProGluAlaThrGlnArgIleThr	219
Db	1026	GCCTCTTATATTTCCCAAACTCATCGCTTTATCAAGCGCAAAACAAAGATCCCGCAC	1085
OY	220	ArgProAspLeuProTyrGlnIleAlaArgArgSerAlaTyrPheSerHisSerHisPro	239
Db	1086	AGGCAGAGATTATATATGAGCCCTCCAAAGAGATCTGGTGGACAAATCATATCTGCACCA	1145
OY	240	ThrGlnSerAlaIleThrProSerSerThrValProLysThrGlnAspGlnArg	259
Db	1146	---CCATCAAAAGCATTCACCA---CTTACCAAGTCCCAAAACGAAAGACCCACAGG	1199
OY	260	ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGly	279
Db	1200	CCACAGCATAGATCCTTATACAGATCTTGGGCAACAAGACACCGACTTGCAAATCCAGGA	1259
OY	280	SerGIuGlnIleGlnLeuThrPheIleLeuGluLeuLeuSerAspSerSerAsnSer	299
Db	1260	AGTGGGCGATTCAGCTCTGGCAATTTCTTATGGAAAGCTTGGAGATGTCACATCC	1319
OY	300	AsnGlyIleThrTrpGluGlyThrAsnGlyGlnPheLysMetThrAspProAspGluVal	319
Db	1320	AACATGCATACTTGGGAAGAACCAATGAGAAATTTAAGATGACCCGATCCCGATGAACTG	1379
OY	320	AlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArg	339
Db	1380	GCAAGACGTTGGGGGAGAGGAAAAAGCAACCCAAATGAGACTTATGACAACTCAGCGGT	1439
OY	340	AlaIleuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAla	359
Db	1440	GCACCTTGCTTACTACTACGATAAAAAATTTATGTACTAAAGTCCATGGCAACGCTATAGCT	1499
OY	360	TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProTyrGlySerSer	379
Db	1500	TACAAATTTGATTTTCATAGGATGGATGGCTTCACACCTCATCTCCAGATTCACAC	1559
OY	380	MethylLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLys	399
Db	1560	ATGTACAAAGTATCCCTCAGAACTTCATATATGTAGCTCATACATGTGCACATCCACAGAG	1619

OY		400	MelanophthalapronhisproproballeuprovalitnrseserSephphala	413
Dd		1620	ATGACCTTGGTCCGCCCATCCGCACCTTACCCTGTGAATGTCAGATTCTTTCA	1679
OY		420	AlaproanpPOTyTTPRasSerProthrGlyGlyIleTyTPrOasnThArGLeuPro	439
Dd		1680	GCCCCATATGATACTGSAATTCACCACTGGAAGTAGTAATTTCACAATAACTCGGCTGCA	1739
OY		440	AlalahisMetProSerHisLeuGlyThrTyTyr	451
Dd		1740	GCTAGCATTATGTCTTCTCATCTTGAACTACTAT	1775
RESULT 11				
AB073078				
LOCUS		AB073078	2131 bp	mRNA linear ROD 17-OCT-2001
DEFINITION				Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,
ACCESSION		AB073078		transcript variant 1, complete cds.
VERSION		AB073078.1	GI:16191716	
KEYWORDS				
SOURCE				
ORGANISM				Mus musculus cdna to mRNA, clone:Erg-1.
REFERENCE				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS				Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE				Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.
JOURNAL				Mus musculus Erg mRNA
REFERENCES				Published Only in Database (2001)
AUTHORS				2 (bases 1 to 2131)
TITLE				Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.
JOURNAL				Direct Submission
REFERENCES				Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel.:81-45-503-9111, Fax:81-45-503-9170)
FEATURES				
SOURCE				Location/Qualifiers
gene				1..2131
CDS				/organism="Mus musculus"
				/db_xref="taxon:10090"
				/chromosome="16"
				/clone="Erg-1"
				1..2131
				/gene="Erg"
				148..1536
				/gene="Erg"
				/note="homolog of human ets-related gene (ERG), transcript variant 1"
				/codon_start=1
				/protein_id="BAB69948.1"
				/db_xref="GI:16191717"
				/translation="MIQTVPDPAHIREALSVSEDSLEFCAVGPRLAKTEMTASSSSDGQSKMSKPRVPOODMLSOAPARYTIKMECPQVNSRNSPDECSYNKGKMYGSPDIVMGSYGYMEKEKHPNPMNTNRBYIVPADPTLMSGNDHYROWLEMAVEYGLLDVDTLFONTDKGLCKMTKDDDRPLRPYSNAOLLSHLYIRTRPIPLHTSDVDKA LONSPRLMARNTDLTYEPBRASAMTHSHLTPOSKAQSPSPAVKTEOORPOLDET QILGPTSSRLANDPGSGIOLOMFLLELLSDSSNNCITWGETNEEFMTDPDEVARM GERKSERNMNADKLRAILRYDYDNIMETKYHGKAYAKFPFHGIAQLAQHPPESSLY KYPDLPYMGSYHAHPQKNMFVSPHPALPVYTSSFPASPBNPYNSPTGIYPTRIDLP ASHMRSLHGYTY"
BASE COUNT		582 a	590 c	530 g 429 t
ORIGIN				
Alignment Scores:				
Pred. No.:		6.05e-141	Length:	2131
Score:		2151.00	Matches:	402
Percent Similarity:		91.07%	Conservative:	16
Best Local Similarity:		87.58%	Mismatch:	27
Query Match:		88.01%	Indels:	14
OB:		10	Gaps:	4

US-09-902-772-2 (1-451) x AB073078 (1-2131)

```

Oy 2 AlaSerThrIleuysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db 172 GCACCTATATATTAAAGAGGAGGCTTGTCAAGTGTGAGCGGAGCAGCTACTATTTGAGAGT 231
Oy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
Db 232 GCCTACGAGACGCCACACCTGGCTAAGACAGAGATGACCGCATCTCTTCCAGTGACTAT 291
Oy 42 GlyIleThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnProPro 61
Db 292 GGCAGACATCCAGAGATGAGTCCAGAGTCCCTCAGAGAGACTGGCTGTCTCAAGCCCA 351
Oy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 352 GCCAGGGTCCACATCAAGATGAGTGCACCCCTAGTCAGGTGAATGTTCCAGCACTCA 411
Oy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 412 CCTATGATGTCAGTGTGACAAAGGTGGAGATGTTGGGACGCCGATACGTGTGGG 471
Oy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThrAsn 121
Db 472 ATGAGCTACGCGACACTACATGAGAGAGAGCATGTGGCGCTCCCAATATGACCAAAAT 531
Oy 122 GluArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArgGln 141
Db 532 GAGGCGAGAGTATGCTCCCTGAGATCTCTCTGAGACACAGACCATGTCCGACAG 591
Oy 142 TyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValaPheLeuLeuPheGln 161
Db 592 TGGCTGAGTGGGGGGGGAAGATATGGCTCTCGATGTGGAGCTTACTATTTTCAG 651
Oy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 652 AATATCGATGGAGAGAGCTGTGCAGATGACAAAGATGACTTCCAGCGGCTCAGCGG 711
Oy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db 712 AGCTACATGCGCATCTCTCTCATCTCCACTACCTCAGACAGACTCCCTTCCA 771
Oy 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro 221
Db 772 CATCTG-----ACTTC-----GATGACGTTGATGAGGCTTTACAAACTCTCCA 816
Oy 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233
Db 817 CGTTAATGATGCGCAGAAACAGATTACCTATGAGCCTCCAGAGATCAGCTCG 876
Oy 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db 877 ACCGCGCACAGCCACCTCAGCTCCTCAAACTCTCAGCATCTCCCTCAGCAGTG 936
Oy 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db 937 CCCAAACTGAAGACAGCGCTCCTCAGTACATCTTACCAAGATCTGGAGCCAGCACT 996
Oy 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGluLeu 292
Db 997 AGCGGCTGCTCTAATCCAGAGTAGGCGCAGATCCAGCTGTGGCAGTTCTCTCGAATC 1056
Oy 293 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheLys 312
Db 1057 CTGTCAACAGACTCCACTCCACTGATCACCCTGGAGAGGACCAACAGGGGAGTTCAAG 1116
Oy 313 MetThrAspProAspGluValAlaArgArgTrpGlyLysGlyLysSerLysProAsnMet 332
Db 1117 ATGACACACCGGAGAGGTGGCTGGCGCTGGGGGAGAGAGACCAAGCCCAATG 1176
Oy 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352

```

```

Db 1177 AACTATGACAGCTACAGCCGCCCTCCGCTACTACTACGACAAACATCATGACCAAG 1236
Oy 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
Db 1237 GTGACAGGGAAGCCCTACGCTACAGATTGACTTCCACAGGAGATGGCCAGCCCTGCAG 1296
Oy 373 ProHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392
Db 1297 CCCACCCCTCTGAGTGTCTCTTACAGATACCCCTCCGACTGCCATATATGAGCTCC 1356
Oy 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db 1357 TATCAGCGCCACCCCAAGAGATGACTTGTGTCTCCCGCCCTCCGCTCTCCAGTCC 1416
Oy 413 ThrSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyLe 432
Db 1417 ACATCTTCCACTTTCTTGTCTCCCGAACCACATGGAATTCACGACATGGGGGATC 1476
Oy 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1477 TACCGAAGACTAGGCTCCAGCAGCATATGCCCTCTCAGCTGGGACCTACTAC 1533

RESULT 12
LOCUS HUMERG11 3126 bp mRNA linear PRI 27-APR-1993
DEFINITION Human erg protein (ets-related gene) mRNA, complete cds.
ACCESSION M21535
VERSION M21535.1 GI:182182
KEYWORDS erg protein.
SEGMENT 1 of 2
SOURCE Human, cell line COLO 320, cDNA to mRNA, lambda-7.
ORANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 3126)
AUTHORS Reddy,E.S., Rao,V.N. and Papas,T.S.
TITLE The erg gene: a human gene related to the ets oncogene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (17), 6131-6135 (1987)
MEDLINE 87317608
PUBMED 3476934

FEATURES
source
1..location/Qualifiers
1..3126
/organism="Homo sapiens"
/db_xref="taxon:9606"
195..1286
/feature="erg1 protein"
/codon_start=1
/protein_id="AA035811.1"
/db_xref="GI:182185"
/translation="MVGSPDTVGNNVSGMEKKHMPNPNTTNERRYIVPADPTLWST
DHYROMLEMAVKEYGLPDVINILFQNDGKEICKMTDDQRLTPSINADILLSHLY
LRETRPLHLTSDVDKALONSPRLMHARKNDLPLEPRKSAWGHGHPROSKAOPS
PSTVYKTEIDRPOLDVQIIGPSSRLASGSGIOIMQFLLELLSDSSNSCTWEG
TNGEFTKTDDEVARLMGERKSRNNMYDKLSRLRYRYKNITKVKGRYAYKPDF
HGIALOLQPHPESSLYKPSDLPYMSYHAHPOKNFVAPHPALPVTSSFFAAN
PYNPSPTGTYIPNTRLPESHMPSHLGTY"
BASE COUNT 928 a 732 c 725 g 741 t
ORIGIN 1 bp upstream from EcoRI site.

Alignment Scores:
Pred. No.: 5.6e-117 Length: 3126
Score: 1810.00 Matches: 336
Percent Similarity: 90.34% Conserved: 10
Best Local Similarity: 87.73% Mismatches: 23
Query Match: 74.06% Indels: 14
Gaps: 4

US-09-902-772-2 (1-451) x HUMERG11 (1-3126)
Oy 78 SerArgAsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSer 97
Db 150 TCAAGAACTCTCTGTGATGATGATGAGTGTGGCCAAAGCGGAGAGATGTGGGACGCCA 209

```

OY	98	AspAsnValGlyMetAsnTyrIleGlySerTyrMetGlnGluLysHisIleProProAsn	117
Db	210	GACACCGTTGGGATGAACTACGGGACGCTACATGAGGAGACACACTGCCCCAAC	269
OY	118	MetThrThrAsnGluArgArgValIleValProIleAspProThrLeuTyrSerThrAsp	137
Db	270	ATGACACGAGAACGACGCGACAGATGTTTCGTCGACGACGATCTTACGGTATGGATACAGAC	329
OY	138	HisValArgGlnTrpLeuGlnTrpAlaValLysGlnTyrGlyLeuProAspValAspIle	157
Db	330	CATGGCGGCACTGGCTGGAGGGCGGGGAAAGATATGGCTTCCACACGTCACATC	389
OY	158	LeuLeuPheGlnAsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGln	177
Db	390	TTCGTTATCCAGAACATCATGATGGAGGAAGACTGTGCAGATGACCAAGACGACTCCAG	449
OY	178	ArgLeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlu	197
Db	450	AGGCTCACCCCGACGTACACGCGACACCTCTCTCACAATCCACTCAGTACCTCAGAGAG	509
OY	198	ArgGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIle	217
Db	510	ACTCCTCTCCACATTTG-----ACTTCA-----GATGATGTTGATTAAGCGTTA	554
OY	218	ThrThrArgPro-----AspLeuProTyrGlnGluAlaArg	229
Db	555	CAAAACTCTCCACGGTTATGATGATGCTAGAAAACAGAGATTTCCATATAGCCCCACAG	614
OY	230	ArgSerAlaTrpThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSer	248
Db	615	AGATCAGCGCTGACCGGTCACGCGCCACCCACGCCCCAGTCCAAAGCTCTCAACATCT	674
OY	249	SerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeu	268
Db	675	CTTCTCCACAGTGCACAAACGTAAACACAGCCTCTCAGTTGATGACTTATCAGATTCTT	734
OY	269	GlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPhe	288
Db	735	GGACCAACAGTAGCGCGCTTGCAATCCAGGACATGGCCACATCCAGCTTTGGCAGATTC	794
OY	289	LeuLeuGlnLeuLeuSerAspSerSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsn	308
Db	795	CTCCTGGAGCTCTGTCGGACAGCTCCAACTCAGCTCATGACCTCGGGAAGACCAAC	854
OY	309	GlyGluPheLysMetThrAspProAspGluValAlaArgGlyTrpGlyGluArgLysSer	328
Db	855	GGGAGTTCCAAAGATCAGGATGCCACACAGAGTGGCGCGCTGGGGAGACGGGAAGAC	914
OY	329	LysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsn	348
Db	915	AAACCCAACTAGAACTAGATAGCTGACCGCGGCTTCCTACTACTATACAAAGAC	974
OY	349	IleMetThrLysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAla	368
Db	975	ATCATGACCAAGATCCATGAGGAGGCGCTACAGGCTTCAAGTTCCAGTCCACGCGAATCGCC	1034
OY	369	GlnAlaLeuGlnProHisProProGlnSerSerMetTyrLysTyrProSerAspLeuPro	388
Db	1035	CAGGCGCTCCAGCCCCACCCCCCGAGTCACTCTGTGTCAATACATCCCTCAGAACTCCCG	1099
OY	389	TyrMetSerSerTyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProPro	408
Db	1095	TACATGGGCTCTATACAGCGCCACCCAGAAAGATGAATCTTGTCGGCCCACTTCA	1155
OY	409	AlaLeuProValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerPro	428
Db	1155	GCCTCTCCCGAGACATCTTCCAGTTTCTTCTGCCCCCAAAACCCCTACTGGAATTCACCA	1214
OY	429	ThrGlyGlyIleTyrProAsnThrArgLeuProAlaIleHisMetProSerHisLeuGly	448
Db	1215	ACTGGGGGTATATACCCCAACACTTAGGCTCCCCACAGCCATATGCTTTCATCTTGGGC	1274
OY	449	ThrTyrTyr 451	

DB	1275	ACTTACTAC	1283		
RESULT 13	DRE249590		1436 bp	mRNA	linear
LOCUS	DRE249590				
DEFINITION	Danio rerio mRNA for Fil-1 protein (fil-1 gene).				
ACCESSION	AJ249590	GI:6006476			
VERSION	AJ249590.1	GI:6006476			
KEYWORDS	fil-1 gene, Fil-1 protein.				
SOURCE	Danio rerio.				
ORGANISM	Danio rerio				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.				
AUTHORS	1 (bases 1 to 1436)				
TITLE	Brown, J.A., Rodaway, A.R., Schilling, T.F., Jowett, T., Ingham, P.W., Patient, R.K., and Sharrocks, A.D.				
TITLE	Insights into early vasculogenesis revealed by expression of the Ets-domain transcription factor Fil-1 in wild-type and mutant zebrafish embryos				
JOURNAL	Mech. Dev. 90 (2), 237-252 (2000)				
MEDLINE	20108585				
PubMed	10640707				
REFERENCE	2 (bases 1 to 1436)				
AUTHORS	Sharrocks, A.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-SEP-1999) Sharrocks A.D., School of Biological Sciences, University of Manchester, Oxford Road, Manchester, M13, 9PT, UNITED KINGDOM				
FEATURES	Location/Qualifiers				
Source	1..1436				
	/organism="Danio rerio"				
	/db_xref="taxon:7955"				
	/dev_stage="20-24hr embryo"				
	45..1400				
	/gene="fil-1"				
	45..1400				
	/gene="fil-1"				
	/function="transcription factor"				
	/codon_start=1				
	/evidence=experimental				
	/product="Fil-1 protein"				
	/protein_id="CAB56832.1"				
	/db_xref="GI:6006477"				
	/translation="MDSITKEALSVSSEDSOLFEPYAAAPLPEKTDWTASGTODYGC				
	TKHINPIPOQEWINQPVNVKREYHINSRESYPDCSVGKCNKVGGEASQMNVT				
	TGYDEKCAPPPNMTTNERAVIYPADSLMSPDHQWLDAIKKEYGLQEDTAMFHS				
	TGDEKMSKDDFLRLTSYNTVELLSHLNLABESSISINYPNSHADOSPRLAARD				
	DASVDARITGCSNMNSMGKSGPSVQSQSVKNDPDRPODPYOIIGPSSRLANGC				
	SGQIOIOWFLLELSDSNACCTMEGTNGEFGKATDPDEVARRMGERKSRNNATDLK				
	SRARITTDKINIKTRVHGKRTATVFDFHGLAQALQPHFTESTIKITPSLPTVPSYHA				
	HQQKVNVSFPPSPMPVYTSNMFEGPTTPYMSPPGILPNPSVPRHANSHVSLGISTY				
gene					
CDS					
ORIGIN	377 a 439 c 349 g 271 t				
ALIGNMENT SCORES:					
Pred. No.:	2.54e-105	Length:	1436		
Score:	1637.00	Matches:	303		
Percent Similarity:	78.90%	Conservative:	56		
Best Local Similarity:	66.59%	Mismatches:	88		
Query Match:	66.98%	Indels:	8		
DB:	5	Gaps:	5		
US-09-902-772-2 (1-451) x DRE249590 (1-1436)					
Oy	1	MetaLaserThrltelysGLuAlaLeuSerValValSerGIuaSPgInSerLeuPheGlu	20		
Db	45	ATGACGACGACATTTAAGGAGCGCCTGTCACTGGTGAAGACCACTCTCTTTGCAG	104		
Oy	21	CysAlaTyr--GlySerProHisLeuAlaLysThrGluMetThrAlaLeuSerSerSer	39		

```

Db 105 CCTCATACGCCGCTGCTCTTATCCCAAGACAGACATGACTGCCGCGGACTCAG 164
Qy 40 GlnTyrLysGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln 59
Db 165 GATTACGGCCAGACGACCAATATCAACCCATTCCTCCCTCAGCAAGAGTGAATTAACAG 224
Qy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 225 ---CCGGTGGCGGGAACGTCACCGAGATAT-----GACCACATCAACGGATCCAGA 275
Qy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 276 GAGTCGCCGGTGCAGCTCAGCGTGGGAAATGCAATAAATGGTGGCGGAGACCGAGCG 335
Qy 100 ValGlyMetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThr 119
Db 336 TCTCAGATGAACTACTGCTGCTACATGATGAGAAAGTCTCCGCTCCCAACATACAG 395
Qy 120 ThrAsnGlnArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisVal 139
Db 396 ACCAATGAGAGAGATCATCGTCCCGCAGACCCGCTCTGTGTGCTCCGATCAGCTG 455
Qy 140 ArgGlnTyrLeuGlnTyrPalValLysGlyLysLeuProAspValAspIleLeu 159
Db 456 CGGAGTGTAGTACGTGGCCATTAAGAGTACGCTCTCAGAGATCGACACGCGCATG 515
Qy 160 PheGlnAsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArgLeu 179
Db 516 TTTTCACGACAGAGTGGAAAAGACTCTGCAGATGAGCAAGCGACTTCCTGAGACTC 575
Qy 180 ThrProSerTyrAsnAlaAspIleLeuSerHisIleAsnHisTyrLeuArgGlyLys 199
Db 576 ACCGAGCTTATTAACACCGAGCTCTCTCATCATCTCAATTAACCTCAGGAAAGTACG 635
Qy 200 AlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThr 219
Db 636 TCATCGCATATCATAC---AACACGCCATCTCAGCTGACACGACCCCTCCGCTGCA 692
Qy 220 ArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaThrThrSerHisSerHisPro 239
Db 693 AAGACATGCGCTTATGATGCTGATGAGGAGGCGGCTGTCGAACAACATGCACAGC 752
Qy 240 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlnAspGlnArg 259
Db 753 GGAAGAAAGCTCTCCAAAGTGTCTCAGAGCGTGTCCAAAACCCGGAATCAGCTCTGT 812
Qy 260 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGly 279
Db 813 CCTCAGCGCATCTCTTATCAGATATTAAGACCCACGAGTGCAGTTCCTTAACCCAGGC 872
Qy 280 SerGlyGlnIleGlnLeuTyrPheLeuLeuGlnLeuLeuSerAspSerSerAsnSer 299
Db 873 TCAAGTCAGATCCAGCTGTGGAGTCTCTGTGAGCTCTGTCCGACCGCCAGCGC 952
Qy 300 AsnCysIleThrTyrPheGlnGlyLysAsnGlyLysPheLysMetThrAspProAspGlyVal 319
Db 933 GGCTGCTACCTACCTGGGAGAGAACCAAGGAGATTCAGATGACGAGCCACGAGAGGTG 992
Qy 320 AlaArgArgTyrPheGlyLysSerLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 339
Db 993 GCACGGCGCTGGGGCGAAGCAAGACCAACCAACATGAATGACGACGAGCTGAGCGGA 1052
Qy 340 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAla 359
Db 1053 GCCCTGGCTACTACTACGACCAAGAACATCATGCAAACTGACGCGCAAGCCCTTACGCC 1112
Qy 360 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisPropGlnSerSer 379
Db 1113 TACAGTTCGACTTCACGCGCATTCGCCAGCCCTGCAACCCATCCACCGAGTCCACC 1172
Qy 380 MetCysLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLys 399
Db 1173 ATGTACAAATACCAAGCGAGCTGCGTACGTGCTATATCAGCCGCCACGACGAGAAG 1232

```

```

Qy 400 MetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAla 419
Db 1233 GTCAACTTCGTCTCTCTCCACCCCTCGTCATATGCGCTGACCTCTCCAACTTCTTGGT 1292
Qy 420 AlaProAsnProTyrTyrPAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuPro 439
Db 1293 CCCCAACCCCATATCTGAGCTGCCCAACGAGAGTATTTACCCCAATCCACAGCTGCCA 1352
Qy 440 Ala-----AlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1353 CGCCATGCCAATTCACAGCTGCTTACACCTTGATGTTACTAC 1397

RESULT 14
AY029368 1359 bp mRNA linear PRI 09-MAY-2001
LOCUS AY029368
DEFINITION Homo sapiens friend leukemia integration 1 transcription factor
ACCESSION AY029368
VERSION AY029368.1 GI:14017402
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1359)
AUTHORS Udhil,B.T.S., Kainey,D.R. and Meredith,D.M.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2001) Molecular Medicine Unit, Leeds University,
Beckett St, Leeds LS9 7TF, United Kingdom

FEATURES
source
1..1359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
1..1359
/gene="FHL1"
1..1359
/gene="FHL1"
/note="FHL1"
/note="FHL1 transcription factor; involved in the
pathogenesis of Ewings sarcoma following translocation
with EWS gene on chromosome 22; may have oncogenic
potential when fused with EWS gene"
/codon_start=1
/product="friend leukemia integration 1 transcription
factor"
/protein_id="AAK50443.1"
/db_xref="GI:14017403"
/translation="MDGTEIKELSVSDQSLPDSATGAHAHLPRADMTASGSPDYGQ
PKRINLPDQEWLNQPVRYNVRREYDHNMGRESPPDCSVSKCSKLYGGESPNHY
NSTADEKNGPPPPNMTTNERRVVAPADPTLMTQDEHROWLEAIKETSLEIDTSTFO
MDGKEIKKMKKEDFLRATLLYNTVELSHLSYLRSSSLAYWTSHTDSSRLSYVE
DPSYDSVRRGAMWNMNSGLNKSPLEGAGTIEKNTKEORQPDYQILGPTSSRLNP
GSGOILMOFLLELLSDSANASCITMEGTGTEKMDPDEVARRMGEKSKPMNYDK
LSRALRYDYKNTMTVYHGKRYAVKDFEHLIAQALDPHPESSMYPSISYSPSVH
AHOOKXNFPVPHRHSMPYVTSRFFGAASQVWTSPTGITYINPVRPHRPNHYVSHLS
Y"

BASE COUNT 359 a 422 c 333 g 245 t
ORIGIN
Alignment Scores:
Pred. No.: 1,77e-104 Length: 1359
Score: 1624.50 Matches: 306
Percent Similarity: 79.43% Conservative: 57
Best Local Similarity: 66.96% Mismatches: 83
Query Match: 66.47% Indels: 11
Gaps: 7
DB: 9

US-09-902-772-2 (1-451) x AY029368 (1-1359)
Qy 1 MetAlaSerThrIleLysGlnAlaLeuSerValValSerGlnAspGlnSerLeuPheGln 20
Db 1 ATGAGCGGAGCTATTAAGAGGCTCTCTGCTGTGAGCGACGACGACGCTCCCTTTGAC 60

```



```

Oy 21 CysAlaTyrGly----SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
Db 61 TCAGCGTACGAGCGGCGAGCCATCTCCCAAGGCGCAGCATGCTCGGGAGTCT 120
Oy 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGln 59
Db 121 GACTACGGGCGAGCCCAAGATCATCCCTCCACACGACGAGGATGATCATCATCAG 180
Oy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 181 ---CCAGTGGAGGTCAACGTCACAGGAGAT-----GACCACATGAAATGATCCAGG 231
Oy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 232 GAGTCTCGGTGACTGAGCGGTAGCAAAATGACAGCAAGCTGTGGGCGGAGGAGTCC 291
Oy 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProProAsnMet 118
Db 292 AACCCCAAGACTACACAGCTATATGAGAGAGATGAGCCCTCCCTCCACACATG 351
Oy 119 ThrThrAsnGluArgArgValIleValProAlaAspProThrIleThrSerThrAspHis 138
Db 352 ACCACCAAGAGAGAGAGATGATGTCCTCCGACAGCCACACTGTGACACAGAGCAT 411
Oy 139 ValArgGlnThrLeuGluThrPalaValLysGluTyrGlyLeuProAspValAspIleLeu 158
Db 412 GTGAGCGAATGGCTGAGTGGCGCATTAAGAGATGACAGCTGATGAGATGACACATCC 471
Oy 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
Db 472 TTTTTCGAGAACAGATGAGCAGAACCTGTGTAATGAAAGAGGAGACTTCCCTCCG 531
Oy 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLysIleTyrLeuArgLys 198
Db 532 GCCACCAACCTCTCAACACAGAGAGTCTGTGTGCACACCTCACTTACCTAGGAGAA--- 588
Oy 199 GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr 218
Db 589 ---AGTTCAGTCTGGCTATATATACAACTCCACACCGCAATCCTCAGATTGAGT 645
Oy 219 ThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaThrPheSerHisSerHis 238
Db 646 GTCAAAGAAAGACCTTCTTATGACTCAGTACAGAGAGAGCTGTGGCGCAATAATCATGAT 705
Oy 239 ProThrGlnSerLysAlaThrGln---ProSerSerSerThrValProLysThrGluAsp 257
Db 706 TCTGGCTCAACAAAGTCTCCCTTGGAGGGGACACAAAGCATCAGTACAGATACAGAG 765
Oy 258 GlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
Db 766 CAAGGCGCCAGCAGATCGTATCGATCATCTGGGCGCCGACACAGATCGGCTGAGCCAAC 825
Oy 278 ProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeuSerSerSerSer 297
Db 826 CCTGGAAGCGGCGAGATCGATCGTGTGCAATTCCTCTCGGCGCTCTCGAGAGCGCC 885
Oy 298 AsnSerAsnCysIleThrTyrPglGluGlyThrAsnGlyLysPheLysMetThrAspProAsp 317
Db 886 AACCCCACTGTATCAGCTGGAGGGAGCCACAGGGAGTTCAAAATGACGAGCCCGAT 945
Oy 318 GluValAlaArgArgTyrPglGlyLysGlySerLysProAsnMetAsnTyrAspLysLeu 337
Db 946 GAGTGGCGCAGGCGCTGGGCGAGCGAAAGCAAGCCCAATGATGATGACAGACAGCTG 1005
Oy 338 SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArg 357
Db 1006 AGCGGGCGCTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
Oy 358 TyrAlaTyrLysPheAspPheHisGlyTleAlaGlnAlaLeuGlnProHisProProGlu 377
Db 1066 TATGCTTCAAAATTTGACTTCCACGCGCATTTGCCAGGCTCTGACGACCATTCGACCGAG 1125

```

```

Oy 378 SerSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHisAlaHisPro 397
Db 1126 TCGTCATGTCACAGTACCTTCTGACATCTCTACATGCTCTTCAATGATGCCACAG 1185
Oy 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerPhe 417
Db 1186 CAGAGGAGGAACTTTGCTCCCTCCATCCATCCATCCATGCTGCTGCTGCTGCTGCTGCT 1245
Oy 418 PheAlaAlaProAsnProTyrTyrPAsnSerProThrIleGlyIleTyrProAsn----- 435
Db 1246 TTTGAGACCGCATCATCATATGAGACTCCCGCCAGGGGGAATCTACCCCAACCCCAAC 1305
Oy 436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1306 GTCCCGCGCATCTTACACCCAGCTGCTTACACACTTACAGCAGTACTACTAC 1356

RESULT 15
BC010115
LOCUS BC010115 2394 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:19589 IMAGE:3635042, mRNA, complete cds.
ACCESSION BC010115
VERSION BC010115.1 GI:14603315
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 2394)
Strausberg, R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
nisc.mgc@hgrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Mastaglio, C., Mastrian, S.D., McCluskey, J.C.,
Mcdowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J.,
Tlongson, E.E., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LMNL at: http://image.lnl.gov
Series: IRAL Plate: 27 Row: 9 Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA, gi: 257353.

FEATURES
source
location/Qualifiers
1..2394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:19589 IMAGE:3635042"
/tissue-type="Lymph. Burkitt lymphoma"
/clone_id="NIH MGC.8"
/lab_host="DH10B-R"
/note="Vector: pOFB7"
126..1484
/codon_start=1
/product="Unknown (protein for MGC:19589)"
/protein_id="AAH0115.1"
/db_xref="GI:14603316"
/translation="MDGTKEALSVSDQSLPDSAYGAANHLPKADMTASGSPYGO

```


PHKINLPPOEEMINOPYRVNKRREYDHMANGSRESYPVDCSVSKSKLGGESNPMNY
 NSYDEKMGPPPMPTMTERERVYIPADPTLMTQHRQMLFMAIKEYSIAMTIDISFRO
 NMDGKEJCKMKNKEDFLRATITTYNTEVLSHSTLEBSLLANTTSHDQSRISVKE
 DPSTDSVRGAMGNMNSGLNKSPPGGAOTISKNTEDRPPDPQIIGPSSRLANP
 GSGOIQIMQFLEELSDSANASCITMETNEDEKMTDDEVARMRGEEKSRPNMYDK
 LSRALRYDDNIMTKYHKNRYAYKFDHNGIAOALQDPTESSMYKYSYSDSYMSYH
 AHQOKVNFVPPSPSMPTSSFFGAAASQWYTSPTGIYIPNPNVPRHNTNHPSHLS
 YH"

BASE COUNT 669 a 609 c 574 g 542 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,47e-104 Length: 2394
 Score: 1624.50 Matches: 306
 Percent Similarity: 79.43% Conservative: 57
 Best Local Similarity: 66.96% Mismatches: 83
 Query Match: 66.47% Indels: 11
 DB: 9 Gaps: 7

US-09-902-772-2 (1-451) x BC010115 (1-2394)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGln 20
 DB 126 ATGAGCGGGACTATTAAGAGGCTCTGCGGTGGAGCGAGCACAGCTCCCTCTTGAC 185
 QY 21 CysAlaIleArgLys---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
 DB 186 TCAGGCTTACGAGCGGCGACGCCATCTCCCAAGGCGGACATGACCTCGGGGAGTCT 245
 QY 40 GluIleArgLysGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGln 59
 DB 246 GACTACGGGCGAGCCCAACAGATCAACCCCTCCACAGCAGAGAGTGGATCAATCAG 305
 QY 60 ProProAlaArgValThrIleLysMetLysAsnProAsnGlnValAsnGlySerArg 79
 DB 306 ---CCAGTGGAGGCTCAACGTCACAGGCGAGTAT-----GACCAATGAAATGATCCAGG 356
 QY 80 AsnSerProAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
 DB 357 GAGTCTCCGGTGGAGCTGACGCTTACCAATGACCAAGCTGGTGGGCGAGCGAGTCC 416
 QY 100 ValGlyMetAsnTyrgLysSerTygMetGluGluLysHis---IleProProAsnMet 118
 DB 417 AACCCCATGAACTCAACAGCTATATGAGAGAGAAATGGCCCCCTCCCTCCCAACTG 476
 QY 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuThrSerThrAspHis 138
 DB 477 ACCACCAAGAGAGAGATCATGCTCCCGACAGCCACACCTGTGACACAGAGCAT 536
 QY 139 ValArgGlnThrPheGluThrAlaValLysGluIleProAspValAspIleLeu 158
 DB 537 GTGAGGCAATGGTGGAGTGGCCATTAAGAGATGACAGCTGATGAGATGACACATCC 596
 QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
 DB 597 TTTTCCAGACATGATGAGCAAGAACTGTAAATAAGACAAGAGAGACTTCCTCCG 656
 QY 179 LeuThrProSerTyrgAsnAlaAspIleLeuLeuSerHisLeuHisTyrgLeuArgGluArg 198
 DB 657 GCCACACACCTCTACACACGAGAGTGTGTGTACACACCTCAGTTACCTCAGGAA--- 713
 QY 199 GlyAlaThrPheIlePheProAsnThrSerValTyrgProGlnAlaThrGlnArgIleThr 218
 DB 714 ---AGTTCAGTGTGGCTTAATTAACCTCCACACACCAATCTCAAGATTTGAGT 770
 QY 219 ThrArgProAspLeuProTyrgLysGluAlaArgArgSerAlaTrpThrSerHisSerHis 238
 DB 771 CTCAAAGAAAGACCTTCTTATGACTCACTCAGAGAGAGAGCTTGGCAATTAACATGAT 830
 QY 239 ProThrGlnSerLysAlaThrGln---ProSerSerSerThrValProLysThrGluAsp 257
 DB 831 TCTGCGCTCAACAAAGTCTCTCCCTTGGAGGCGACACAAAGATCAGTAAGATAACAGAG 890

QY 258 GlnArgProGlnLeuAspProTyrgGlnIleLeuGluIleProThrSerSerArgLeuAlaAsn 277
 DB 831 CACGCGCCCGACCGACGATCCGATACGATCTCTGGCCCGGACGACGAGTCCCTACCCAAAC 950
 QY 278 ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSer 297
 DB 951 COTGGAACCGGCGACATCCACTGCGCAATTCCTCCGAGCTCTCTCCGACAGCGCC 1010
 QY 298 AsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMetThrAspProAsp 317
 DB 1011 AACGCCACCTGTATCACCCTGGAGGAGGACCAAGGAGGAGTCAAAATACAGGACCCGAT 1070
 QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrgAspLysLeu 337
 DB 1071 GAGGTGGCCAGCGCTGGGGGAGGAGGAGAAAGCAAGCCCAATGAATTAAGACAAAGCTG 1130
 QY 338 SerArgAlaLeuArgTyrgTyrgTyrgAspLysAsnIleMetThrLysValHisGlyLysArg 357
 DB 1131 AGCGGCGCCCTCCCTTATTAATGATATAAACAATTATAGCAAAAGTCCAGGCAAAAGA 1190
 QY 358 TyrAlaTyrgLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
 DB 1191 TATGCTTCAAAATTTGACTTCCACGGCAATTCGCCAGGCTCTGACGCCATCCGACGAG 1250
 QY 378 SerSerMetTyrgLysTyrgProSerAspLeuProTyrgMetSerSerTyrgHisAlaHisPro 397
 DB 1251 TCGTCCATGTACAAATGACCTTCTGACATCTCTCCATCATGCTCTTACCATGCCCCACAG 1310
 QY 398 GlnLysMetAsnPheValAlaProHisProAlaLeuProValThrSerSerSerPhe 417
 DB 1311 CAGAAAGGGAATTTGTCCCTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 1370
 QY 418 PheAlaAlaProAsnProTyrgTrpAsnSerProThrArgLysIleTyrgProAsn----- 435
 DB 1371 TTTGGAGCGGATCAATTAATGACCTGACCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1430
 QY 436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrgTyrg 451
 DB 1431 GTCCCGCCCATCTCAACACCCACGCTGCTTACACCTTAAGCAGCTACTAC 1481

Search completed: July 28, 2003, 06:34:10
 Job time : 3619.89 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

SUMMARIES

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:05:39 ; Search time 289.339 Seconds

(without alignments)
3510.246 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444
Sequence: 1 MASTKEALSVSEDSLEF.....IYENTRLPAHMSHIGTY 451

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh
-O=/c9r2_1/USPTO.spool/US09902772/runal_23072003_093656_14868/app_query.fasta_1.1294
-DB=N.Geneseq_101002 -QEMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.ccl
-LIST=45 -DOCALLIG=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -GLIC=15
-MOB=LOCAL -OUTPRT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09902772 -CGEN_1_1_364 -trunal_23072003_093656_14868 -NCPU=6 -ICPU=3
-NO_MMAP -PARSEQERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : N.Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	2444	100.0	1447	20	AA26551	DNA encoding chick
2	2414.5	98.8	1528	20	AA26552	DNA encoding chick
3	1624.5	66.5	2938	14	AA050644	Human Hum-F11-1 ge
4	1624.5	66.5	2957	24	ABR84139	Human cDNA differe
5	1599	65.4	2954	22	AAH02915	Human shear stress
6	1506.5	61.6	1890	14	AA050662	Human Hum-F11-1 ge
7	948	38.8	567	22	ABA056124	Human breast cell
8	948	38.8	567	22	ABA66003	Human foetal liver
9	948	38.8	567	22	ABA33090	Probe #11556 for g
10	948	38.8	567	22	AAK14426	Human brain expres
11	948	38.8	567	22	AAK40160	Human bone marrow
12	948	38.8	567	22	AAI20932	Probe #10865 for g
13	948	38.8	567	22	AAI46176	Probe #14862 used
14	948	38.8	567	22	AAI06643	Probe #6634 used t
15	948	38.8	567	24	AB514215	Human genome-deriv
16	863	35.3	473	22	ABA50193	Human breast cell
17	863	35.3	473	22	ABA68128	Human foetal liver
18	863	35.3	473	22	ABA35152	Probe #13618 for g
19	863	35.3	473	22	AAK16512	Human brain expres
20	863	35.3	473	22	AAK42265	Human bone marrow
21	863	35.3	473	22	AAI48359	Probe #17025 used
22	863	35.3	473	22	AAI08696	Probe #8687 used t
23	712	29.1	549	24	ABR44114	CDNA #54 encoding
24	654	26.8	420	21	AAOC1342	Human secreted pro
25	650	26.6	420	22	ABA45056	Human breast cell
26	650	26.6	420	22	ABA55529	Human foetal liver
27	650	26.6	420	22	ABA25239	Probe #3705 for ge
28	650	26.6	420	22	AAK03760	Human brain expres
29	650	26.6	420	22	AAK29224	Human bone marrow
30	650	26.6	420	22	AAI35180	Probe #3866 used t
31	650	26.6	420	22	AAI03692	Probe #3683 used t
32	650	26.6	454	22	ABA42995	Human breast cell
33	650	26.6	454	22	ABA53411	Human foetal liver
34	650	26.6	454	22	ABA53187	Probe #1651 for ge
35	650	26.6	454	22	AAK01681	Human brain expres
36	650	26.6	454	22	AAK27129	Human bone marrow
37	650	26.6	454	22	AAI11718	Probe #1651 for ge
38	650	26.6	454	22	AAI33025	Probe #1711 used t
39	650	26.6	454	22	AAI01646	Probe #1637 used t
40	650	26.6	454	24	ABS01685	Human genome-deriv
41	599.5	24.5	2025	23	ABU05371	Drosophila melanog
42	562.5	23.0	899	23	ABU12097	Drosophila melanog
43	511	20.9	1752	22	AAI53672	CDNA encoding Ratt
44	511	20.9	1752	24	AAI38753	Rat lambda73 CDNA
45	498	20.4	318	21	AAI21875	Human secreted pro

ALIGNMENTS

RESULT 1
AA26551
ID AAX26551 standard; DNA; 1447 BP.

AC AAX26551:
XX
XX 14-JUN-1999 (first entry)

DE DNA encoding chicken C-11 protein.

XX Chicken: C-11 protein; cell calcification inhibiting activity;

KW cell calcification inhibiting agent; c-ery protein; arthritis deformans;

KM ossification; spinal column ligament; ss.

XX Gallus sp.

OS
FH Key Location/Qualifiers
FT CDS 63..1418

```

FT      /tag- a
XX      JP11075871-A.
XX      23-MAR-1999.
XX      29-MAY-1998; 98JP-0166076.
XX      20-JUN-1997; 97US-0050297.
XX      18-JUN-1997; 97US-0878177.
XX      (CHUS ) CHUGAI PHARM CO LTD.
XX      (TYPE-) UNIV PENNSYLVANIA.
XX      WPI: 1999-257708/22.
XX      P-PSDB: AAY01520.
XX      An active protein for inhibiting cell calcification - useful for
XX      measuring the calcification of a cell, for diagnosing arthritis
XX      deformans or ossification of spinal column ligament
XX      Disclosure; Page 7-8; 15pp; Japanese.
XX      The present sequence encodes a chicken C-11 protein which has cell
XX      calcification inhibiting activity. The specification also describes
XX      a cell calcification inhibiting agent containing c-eryg protein
XX      (AAY01521). The proteins are used for measuring the calcification of a
XX      cell, for diagnosing arthritis deformans or ossification of spinal column
XX      ligament.
XX      Sequence 1447 BP; 440 A; 374 C; 317 G; 316 T; 0 other:
SO      Alignment Scores:
Pred. No.: 2,58e-197 Length: 1447
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-902-772-2 (1-451) x AAX26551 (1-1447)
QY      1 MetLaserThrIleIysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      63 ATGGCAAGCACTATTAAAGAGCAATTATCAGTGTGATGAGAACCCAGTCTTGTGAG 122
QY      21 CysAlaIleTyrGlySerProHisIleuAlaIysThrGluMetThrAlaSerSerSerGlu 40
DB      123 TGTGCTACGAGATGCCGCCACCTTGCAGAAAGACAGAAATGACAGCCTCTTCCAGTGA 182
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
DB      183 TATGGCGCAACATCAAGATGAGAGCCCGCGTTCGCCAGCAGAGCTGTTATCCACAGCC 242
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      243 CCGGCCGAGATTACCATTAAGATGAGATGATCAACCAACGATTAATGGTCAAGGAAT 302
QY      81 SerProAspAspCysSerValAlaIysGlyIlyMetValSerSerSerAspAspVal 100
DB      303 TCACCTATGACTGACCGCTGCGCAAAAGAGGAAATGGTTAGCAATTCAGCAATGTT 362
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThr 120
DB      363 GGGATGAACATATGGAACCTACATGAGAGAGAGATATTCGCGCTCAAAATATGACAAAC 422
QY      121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140
DB      423 AATGAAAGAGAGATTATTTGCCAGCAGATCTTACGTTATGAGACAGACACATGTAAG 482
QY      141 GlnTyrLeuGluTyrProAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
DB      483 CAGTGGCTGAGTGGGAGTGAAGAGATGATGCTTCCAGACGTTGACATCTGTGTTTC 542

```

```

QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB      543 CAGAACATGTATGGGAAAGAGTGTGTAAATATGACCAAGATGACTTCCAGAGACTCAGC 602
QY      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArgGlyAla 200
DB      603 CCGAGCTATACCGAGATATCTCCCTGTCACCTACCTACCTCAGAGAGAGAGAGGCC 662
QY      201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArg 220
DB      663 ACTTTTATTTTCCAAATATCATCATGTTTACCAGAGACCAACGCAAGAAATTAACACAGG 722
QY      221 ProAspLeuProTyrGluGlnAlaIysArgSerAlaTyrThrSerHisSerProThr 240
DB      723 CCAGATTATACCTTATGAGCAAGCAGAGATGACGCTGAGAGATCAGACCTCCACCTCCACT 782
QY      241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260
DB      783 CAGTCAAAACCTACCCACCATCATCTTCAACAGTGCACAAACAGAAACAGACAGCTCTCT 842
QY      261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
DB      843 CAGTTATATCTTATCATGATTTCTTGACGACGACGACGCGCTTGCAATTCAGGAGT 902
QY      281 GlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsn 300
DB      903 GGGCAGATACAGCATATGAGAGTTCCTACTGAGAGCTTCTGCGGACAGCTCCAACTCCAAC 962
QY      301 CysIleThrTrpGluGluTyrAsnGlyGluPheLysMetThrAspProAspGluValAla 320
DB      963 TGCATCACCTGGGAGGCGCAAAATGGGAGTTCAGATGACAGACCTTGATGAAGTGGCT 1022
QY      321 ArgArgTyrPylGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340
DB      1023 CCGGCTTGGGAGAGAGAGAAAGCAACCTAATGATGACATGACAAACTCAGCGCTGCA 1082
QY      341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360
DB      1083 CTTCGCTACTACTATGACAAAAATATATGACTTAAGTTTCATGCTTAACCTTATGCCCTAC 1142
QY      361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerMet 380
DB      1143 AAATTTGATTTTCCAGGAATGCTCAGGCGCTCCAGCTTCACTCCAGATATATCCATG 1202
QY      381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400
DB      1203 TACAAATACCCATCAGACCTCCCTCATATGATGATTCATGACATGACACCCAGAAAGATG 1262
QY      401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAlaAla 420
DB      1263 AACTTGTACCTCCCATCCCTCCCTGCTTCCCGTAACTTCATCCACACTTTTGGCTGCC 1322
QY      421 ProAsnProTyrTyrPAsnSerProThrGlyIleTyrProAsnThrArgLeuProAla 440
DB      1323 CCTATCATACTAGGAGATTCACCACTGGAGGACATCTACCCCAATACAGAGGCTGCCAGCT 1382
QY      441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451
DB      1383 GCTCATATGCTTCCCATCTTGGCAGCTACTATC 1415

```

RESULT 2
 AAX26552
 ID AAX26552 standard; DNA: 1528 BP.
 AC AAX26552;
 XX 14-JUN-1999 (first entry)
 DT
 XX DNA encoding chicken c-ery protein.
 DE
 KW Chicken; C-11 protein; cell calcification inhibiting activity;
 cell calcification inhibiting agent; c-ery protein; arthritis deformans;

KW ossification; spinal column ligament; ss.
 XX Gallus sp.
 OS
 FH Key Location/Qualifiers
 FT CDS 63..1499
 FT /*tag= a
 JPL1075871-A.
 XX
 XX
 XX 23-MAR-1999.
 XX
 XX 29-MAY-1998; 98JP-0166076.
 XX
 XX 20-JUN-1997; 97US-0050297.
 PR 18-JUN-1997; 97US-0878177.
 XX
 XX (CHUGAI PHARM CO LTD.
 PA (UYE-) UNIV PENNSYLVANIA.
 XX
 DR WPI: 1999-257708/22.
 DR P-PSDB: AA01521.
 XX
 XX An active protein for inhibiting cell calcification - useful for
 PT measuring the calcification of a cell, for diagnosing arthritis
 PT deformans or ossification of spinal column ligament
 XX
 XX Disclosure: Page 8-9; 15pp; Japanese.
 XX
 XX The present sequence encodes a chicken c-ery protein. The specification
 CC also describes a chicken C-11 protein (AA01520) which has cell
 CC calcification inhibiting activity and a cell calcification inhibiting
 CC agent containing c-ery protein. The proteins are used for measuring the
 CC calcification of a cell, for diagnosing arthritis deformans or
 CC ossification of spinal column ligament.
 XX
 XX Sequence 1528 BP; 464 A; 395 C; 330 G; 339 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 8,71e-195 Length: 1528
 Score: 2414.50 Matches: 450
 Percent Similarity: 94.14% Conservative: 0
 Best Local Similarity: 94.14% Mismatches: 1
 Query Match: 98.79% Indels: 27
 DB: 20 Gaps: 1
 US-09-902-772-2 (1-451) x AA026552 (1-1528)
 QY 1 MetAlaSerThrIleLeuGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20
 DB 63 ATGGCAACGACTATTAAAGAAAGCATTTACAGTGGAGTGAAGACCACTCCCTGTTTGA 122
 QY 21 CysAlaTyrglySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 123 TGTGCTAGCGATGCGCCACCTTGCAAGAGCAAGAAATGACAGCCTCTCTTCCAGTAA 182
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTIPLeuSerGlnPro 60
 DB 183 TATGGGCAAAACATCAAAAGATAGCCCGCGCTCCCGCAGAGACTGGTTTTCACAGCCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 243 CCGGCGACAGTTTACCATTAATGAGAGTGAAGCCAAACAGTTTAATGGGCAAGGAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 303 TCACCTGATGACTGACGCGTGGCAAAAGAGGAAATAGTTAGACAGTTACAGACAATGT 362
 QY 101 GlyMetAsnTyrglySerTyrMetGlnGlnLysHisIleProProProAsnMetThrThr 120
 DB 363 GGGATGAATATGGAAGCTACATGGAAGAGAAAGATATTCGCTCCAAATATGCAACCC 422
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140

DB 423 AATGAACGAGAGTATTGTGCGACAGATCTCTAGCTATTGAGGACAGACCATGTACCG 482
 QY 141 GlnTPLeuGluTPAlaValLysGlyLeuProAspValAspIleLeuLeuPhe 160
 DB 483 CAGTGGCTGGAGTGGGAGTGAAGAGTATGCTTCCAGAGTGGACATCTGTGTGTTTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 543 CAGAACATTGATGGAAAGAGTTGTGTAATAAGCAAAAGATGACTTCCAGAGCTACAG 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgLysGly----- 198
 DB 603 CCGACCTTAACAGCAGATATCTCTCTGACACCTTACACTTACAGAGAGCTCTCTT 662
 QY 198 ----- 198
 DB 663 CCACATTTGACTTTCAGATGATGTGATGAAGCCTTACAAAACCTCTCCAGGTTAATGCAT 722
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213
 DB 723 GCTAGAAACACAGAGAGGCCACTTTTATTTTCCAAATACATCAGTTTACCCAGAGCA 782
 QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrp 233
 DB 783 ACGCAAGAAATACACACAGGCGACATTTACTTATGAGCAACGAGGAGATCAGCGTGG 842
 QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValPro 253
 DB 843 ACGAGTCAACAGCGATCCCATCTGCAAAAGTACCAACATCATCTTCAACAGATGGCC 902
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
 DB 903 AAAAGAGAGACACAGCTCTCAATGATCTTATCAGATTTCTTGAGCCGACGACGAC 962
 QY 274 ArgLeuAlaAsnProGlnSerGlyLysIleGlnLeuTyrPGLnPheLeuLeuLeuLeu 293
 DB 963 CGCTTTCGAATCCAGGAGAGTGGGCAAGTACAGCATGTGCACTTCTGAGAGCTTCTG 1022
 QY 294 SerAspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyLysLysMet 313
 DB 1023 TCGGACAGCTCCAACTCCAACTGCACTGAGGAGGCAAAATGGGAGTTCAAGATG 1082
 QY 314 ThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsn 333
 DB 1083 ACAGACCTGATGAAGTGGCTCGGCTGGGAGAGAGGAAAGCAAACTTAACATGAAC 1142
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
 DB 1143 TATGCAAACTCAGCGCGTGCACCTTGCCTACTACTATGACAAAATAATATATGACTAA 1202
 QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
 DB 1203 CATGGTAAACCTGATGCTTACAAATTTATTTCCACGGAATCGCTCAGGCCCTCAGGCT 1262
 QY 374 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
 DB 1263 CACCTCCAGAAATATCATGTACATAAATACCATCAGACCTCCCTCATCATAGTTCTCTAC 1322
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
 DB 1323 CATGCAACCGCCCGAAGATACACTTGTACTCCCGCATCCCGCTTCCCGCTTAAC 1382
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 433
 DB 1383 TCATCCAGCTTTTGTGCGCCCTAATATCATAGTGAATTCACCAACGAGGAGCATCTAC 1442
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
 DB 1443 CCCAATACAGAGTGGCCAGCTGCTCATATAGCTTCCCATCTTGGACCTACTAC 1496
 RESULT 3
 AA050644

ID AA050644 standard; cDNA; 2938 BP.
 XX AA050644;
 AC
 XX
 XX 26-MAY-1994 (first entry)
 DT
 XX
 DE Human Hum-Fli-1 gene clone BM025.
 XX
 KW chromosomal translocation; chimeric; chimeric; Ewing sarcoma;
 KW Ews gene; malignant melanoma; hum-fli-1;
 KW primitive peripheral neuroectodermal tumour; human chromosome 11;
 KW human chromosome 22; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 143..1501
 FT /*tag= a
 FT /product= HUM-Fli-1
 FT polyA_signal 2908..2913
 FT /*tag= b
 FT
 XX MO9323549-A.
 XX
 XX 25-NOV-1993.
 PD
 XX 19-MAY-1993; 93WO-FR00494.
 PF
 XX 20-MAY-1992; 92FR-0006123.
 PR
 XX (CNRS) CNRS CENT NAT RECH SCL.
 PA
 XX Aurias A, Delattre O, Desmaza C, Melot T, Peter M;
 PI Plougastel B, Thomas G, Zucman J;
 P1
 XX WPI: 1993-386580/48.
 DR P-PSDB; AAR44556.
 XX
 XX New nucleic acid of EWS gene and its hybrid(s) - contg. gene
 PT sequence involved in chromosomal trans-location, also derived
 PT mRNA, probes, fusion proteins etc., for diagnosis and treatment
 PT of Ewing sarcoma and melanoma
 XX
 PS Disclosure; Fig 7; 123pp; French.
 XX
 CC The probe 11KR1 was used to screen a human marrow cDNA library
 CC (Clontech cat.# HL1058). The clone BM025 was identified and
 CC sequenced. It represents the entire coding region together with
 CC 5'- and 3'-UTRs of the Hum-Fli-1 gene.
 XX
 SQ Sequence 2938 BP; 847 A; 692 C; 675 G; 724 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1.52e-127 Length: 2938
 Score: 1624.50 Matches: 306
 Percent Similarity: 79.43% Conservative: 57
 Best Local Similarity: 66.96% Mismatches: 83
 Query Match: 66.47% Indels: 11
 DB: 14 Gaps: 7
 US-09-902-772-2 (1-451) x AA050644 (1-2938)
 QY 1 MetaLaserThrlletysGUAAlaLeuSerValValSerGluaspInserLeuPheGlu 20
 DB 143 ATGAGCGGACTATTAGAGGCTCTGTGGTGTGAGCGAGCAGTCCTCTTGAC 202
 QY 21 CysAlaTyrGly---SerProHisLeuAlaLysThrLumethrAlaSerSerSer 39
 DB 203 TCAAGCGTAGCGAGCGAGCGCCATCTCCCAAGCGCAGACATGAGCTCGGGAGTCC 262
 QY 40 GUAATGATGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
 DB 263 GACTACGGGACGCCCCACAGATACACCCCTCCACACGACGAGGATGATCATAC 322

QY 60 ProProAlaArgValThrLleLysMetGluCysAsnProAsnGlnLysArg 79
 DB 323 ---CCAGTGAAGGCTCAACGTCACACGGGAGTAT-----GACCACATGAATGATCCAGG 373
 QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
 DB 374 GAGTCTCCGTTGCTGAGTGGCCATTAAGCAATGACAGCAAGCTGTGGCGGAGCGAGTCC 433
 QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---LleProProAsnMet 118
 DB 434 AACCCCATGAACTCAACAGCTATATGAGAGAGAGATAGGCCCTCCCTCCACATG 493
 QY 119 ThrThrAsnGluArgArgValLleValProAlaAspProThrLeuTrpSerThrAspHis 138
 DB 494 ACCACCAAGCAGAGAGAGATCATCTCCCGACAGCCACACATGTCGACACAGAGCAT 553
 QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluLysLysLeuProAspValAspLleLeu 158
 DB 554 GTGAGGCAATGGCTGAGTGGCCATTAAGAGATATAGCTTATGATGAGATCGACATCC 613
 QY 159 LeuPheGlnAsnLleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArg 178
 DB 614 TTTTCCAGAAACATGATGCGCAAGACTGTGTAATGAACAAGAGACTTCCTCCG 673
 QY 179 LeuThrProSerTyrAsnAlaAspLleLeuLeuSerHisLeuHisTyrLeuArgLysArg 198
 DB 674 GCCACACACCTCTACACAGAGAGTCTGTGTCACACCTCATCTACAGAGAA--- 730
 QY 199 GlyAlaThrPheLlePheProAsnThrSerValTyrProGlnAlaThrGlnAlaThr 218
 DB 731 ---AGTTCAGTGGCCCTTATATACACCTCCCTTGAGAGGCGCACAAAGCATAGATACAG 787
 QY 219 ThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThrSerHis 238
 DB 788 GTCAAGAGAACCTCTTATGACTAGTCAGAAAGAGAGAGATGGCAATACATGAT 847
 QY 239 ProThrLysSerLysAlaThrGln---ProSerSerSerThrValProLysThrGluAsp 257
 DB 848 TCTGCGCTCAACAAACATCTCCCTTGAGAGGCGCACAAAGCATAGATACAGAG 907
 QY 258 GlnArgProGlnLeuAspProTyrGlnLleLeuGlyProThrSerSerArgLeuAlaAsn 277
 DB 908 CAACGGCCCAAGCCAGATCCGATACATCTGTGGCCGAGCAGACATCGCTAGCCAC 967
 QY 278 ProGlySerGlyGlnLleGlnLeuTrpGlnPheLeuLeuLysLeuSerSerSer 297
 DB 968 CTGTGAGCGGCGAGATCCAGCTGTGCAATCTCTCTGAGCTGCTCCGACAGCGCC 1027
 QY 298 AsnSerAsnCysLleThrTrpGluGlyThrAsnGlyLysPheLysMetThrAspProAsp 317
 DB 1028 AACGCCAGCTGTATACCTGAGGAGGAGCAGCGGAGCTTCAAAATGAGCGACCCCGAT 1087
 QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeu 337
 DB 1088 GAGGTGGCCAGCGCTGGGCGAGCGGAAAGCAGCCCAACATGATTAACGCAAGCTG 1147
 QY 338 SerArgAlaLeuArgTyrTyrTyrAspLysAsnLleMetThrLysValHisGlyLysArg 357
 DB 1148 AGCGGGCCCTCCCTTATCTATGATGAATAAACAATATATGACCAAGTGCAGCGCAAAAGA 1207
 QY 358 TyrAlaTyrLysPheAspPheHisGlyLleAlaGlnAlaLeuGlnProHisProGlu 377
 DB 1208 TATGCTTACAAATTTGACTTCCACAGCATGGCCAGCTGTGACGACATCCGACGAG 1267
 QY 378 SerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisPro 397
 DB 1268 TCGTCATGTACAAAGTACCTTGTGATCTCTTACATGCTCTTCAACATGCGCCAG 1327
 QY 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe 417
 DB 1328 CAGAGGTGAACATTGTTCCTCCCATCATCTCTCATGCTCTCTCAGCTTC 1387

Oy	41	Phalaebala	probsm	probyr	trrps	nsnsr	pro	fhng	ylg	ylly	etyr	pro	asn	-----	435		
Db	1388	FTTGAGACG	CGACATCA	AAATCTG	AGACTCC	CCCCACG	GGGGG	GGGAA	ATCTAC	CCCAAC	CCCAAC	144					
Oy	436	---	ThArgleu	pro	Ala	Ia	His	met	Pro	Ser	His	Ie	auGly	Thr	Tyr	Tyr	451
Db	1448	GTC	CCCCCG	CGATCT	CTTAC	ACAC	CCAC	CGAT	GCTT	ACACT	TAG	CGAC	GCT	ACT	TAC	1498	
		RESULT 4															
ID	ABK84139	standard;	cdNA;	2957	BP												
XX	ABK84139;																
DT	14-AUG-2002	(first entry)															
DE	Human	cdNA	differentially	expressed	in	granulocytic	cells	#10.									
XX	Human:	ss;	granulocytic	cell;	DNA	chip;	bacterial	infection;									
KW	viral	infection;	parasitic	infection;	protozoal	infection;											
KW	fungal	infection;	sterile	inflammatory	disease;	psoriasis;											
KW	rheumatoid	arthritis;	glomerulonephritis;	asthma;	thrombosis;												
KW	cardiac	reperfusion	injury;	renal	reperfusion	injury;	ARDS;										
KW	adult	respiratory	distress	syndrome;	inflammatory	bowel	disease;										
KW	Crohn's	disease;	ulcerative	colitis;	periodontal	disease;											
KW	granulocyte	activation;	chronic	inflammation;	allergy.												
XX	Homo	sapiens.															
OS	XX																
PN	WO2002	28999-A2.															
PD	11-APR-2002.																
PF	03-OCT-2001;	2001WO-US30821.															
PR	03-OCT-2000;	2000US-237189P.															
XX	(GENE-)	GENE LOGIC INC.															
PA	Beazer-Barclay	Y, Weissman	SM, Yamaga	S, Vockley	J;												
PI	WPI;	2002-435328/46.															
XX	Detecting	granulocyte	activation	by	detecting	differential	expression										
PT	of	genes	associated	with	granulocyte	activation,	which	serves	as								
PT	diagnostic	markers	that	is	useful	for	monitoring	disease	states	and							
PT	drug	toxicity															
PS	Claim	1;	SEQ	ID	No	710;	114pp;	English.									
XX	The	invention	relates	to	detecting	(M1)	granulocyte	(GC)	activation								
CC	(GCA),	by	detecting	the	level	of	expression	of	gene(s)	(Gs)	identified	by					
CC	DNA	chip	analysis	as	given	in	the	specification,	and	comparing							
CC	the	expression	level	to	an	expression	level	in	an	unactivated							
CC	GC,	where	differential	expression	of	Gs	is	indicative	of	GCA.							
CC	Also	included	are	modulating	(M2)	GA	by	contacting	GC	with	an	agent					
CC	that	alters	the	expression	of	at	least	one	gene	in	Gs;	(2)	screening	(M3)			
CC	for	an	agent	capable	of	mod											

[illegible]

OY	239	ProthrIngsrLySLAlaThrGln---ProSerSerThrValProLySThrGluAsp	257
Db	878	TCGGGCCCAACAAAGTCCCTCCCTTGAGAGGGGCAACAAAGATCAGTAAAGAAATACACAG	937
OY	258	GlnATgProGlnLeuAspProTyrglnIleLeuGlyProThrSerSerArgLeuAlaSn	277
Db	938	CAACGGCCCCAGCCAGATCCGTATACAGATCTCGGAGCCGACACAGACAGTGGCTTAGCCAAAC	997
OY	278	ProGlySerGlyGlnIleGlnLeuThrGlnPheIleuLeuGluLeuLeuSerAspSerSer	297
Db	998	CTGGAAGCGGGCAGATCCACCTGTGGCAATTCCTCTGGAGGTGCTCTCGAAGAGGCC	1057
OY	298	AsnSerAsnCySIIeThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAsp	317
Db	1058	AACGCCAGCTTATACCTCTGGAGAGGAGCAACAGGGAGTTCMAAATAGACGACCCCAT	1117
OY	318	GluValAlaIaArgATrpglyGluArgLysSerLysProAsnMetAsnTyrrAspLysLeu	337
Db	1118	GAGGTGGCCAGCGCCTGGGGGAGACGGAAAGCAAGCCCAATGAATTAAGCAACAGCTG	1177
OY	338	SerArgAlaIeuArgTyrrTyrrTyrrAspLysAsnIleMetThrLysValHisGlyLysArg	357
Db	1178	AGCCGGGGCCCTCCCTTATTAATGATATAAAACATTAAGACAAAGTGCACGGCAAAAGA	1237
OY	358	TyrrAlaTyrrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu	377
Db	1238	TATGCTTACAAATTTGACTTCCACAGGCATTCGCCAGGCTCGCAGCCACATCCGACCGAG	1297
OY	378	SerSerMetTyrrLysTyrrProSerAspLeuProTyrrMetSerSerTyrrHisAlaHisPro	397
Db	1298	TCGTCCATGTACAGTACCTCTTGACATCTCTTAACATGCTCTCTACATGGCCACACAG	1357
OY	398	GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhe	417
Db	1358	CAGAGGGAAGCTTGTCCTCCCTCCCATCCATCCATGCTGTACCTTCCCAAGCTTC	1417
OY	418	PheAlaIaProAsnProTyrrTyrrAsnSerProThrGlyGlyIleTyrrProAsn-----	435
Db	1418	TTTGGAGCGGATGCACATACACTGAGACCTCCGCCACGGGGGGAATCTACCCAAACCCCAAC	1477
OY	436	---ThrArgLeuProAlaIaHisMetProSerHisIeuGlyThrTyrr 451	
Db	1478	GTCCCCGCCATCCTTAACACCCACAGCTGCTTACACTTAGGACACTACTACTAC	1528
RESULT 5			
AAH02915			
ID	AAH02915	standard; DNA; 2954 BP.	
XX	AAH02915;		
XX	AC		
XX	AAH02915;		
XX	15-JUN-2001	(first entry)	
XX	DE	Human shear stress-response coding sequence SEQ ID NO: 83.	
XX	KW	Human; shear stress-response protein; vascular disease;	
XX	KW	arteriosclerosis; ds.	
XX	OS	Homo sapiens.	
XX	WO200125427-A1.		
XX	12-APR-2001.		
XX	PF	02-OCT-2000; 2000WO-JP06840.	
XX	ER	01-OCT-1999; 99JP-0280976.	
XX	PA	(KYOW) KYOWA HAKKO KOGYO KK.	
XX	PA	(NOJI/) NOJIMA H.	
XX	PI	Nojima H, Yoshisue H, Odayashi M, Ota T, Kawabata A, Sakurada K;	
XX	PI	Kuga T, Sekine S, Nakamura Y, Sugano S;	

xx	WPI: 2001-266308/27.			
DR	P-PSDB: AAB90792.			
PT	DNA sequences, proteins encoded by them and antibodies against them			
PR	useful in diagnosis and treatment of vascular disease caused by			
PP	arteriosclerosis -			
xx				
xx	Claim 20; Page 462-466; 678pp; Japanese.			
CC	The present invention provides the protein and coding sequences of a			
CC	number of human shear stress response proteins. These are useful in the			
CC	diagnosis, treatment and screening of vascular diseases caused by			
CC	arteriosclerosis, including heart failure, post-PCIA restenosis and			
CC	hypertension.			
xx				
SQ	Sequence 2954 BP; 846 A; 698 C; 683 G; 727 T; 0 other:			
 Alignment Scores:				
Pred. No.:	2.2e-125	Length:	2954	
Score:	1599.00	Matches:	304	
Percent Similarity:	78.99%	Conservative:	57	
Best Local Similarity:	66.52%	Mismatches:	84	
Query Match:	65.43%	Indels:	12	
DB:	22	Gaps:	8	
 US-09-902-772-2 (1-451) x AAH02915 (1-2954)				
OY	1 Metalaserthrileysglunlaaleuservalvalsergluaapnglnsertleuphegclu	20		
Db	173 ATGACGGCGCATTTAAAGAGCGCTGTGGTGGAGCAGCAAGCACATCCCTTTTGAC	222		
OY	21 CysAlatryelyl---serProhlsueualalystrhgumetnhrAlaserSerSer	39		
Db	233 TCAGCGTAGAGGAGCGGAGGCCATCCCCAGCGGCAGCATGATGCTCGGGAGTCCT	292		
OY	40 GlutylglylnThrserlysmetserrProarvalProglnglnasPTripLeuSerGln	59		
Db	293 GACTACGGGGAGCCCCACAAGATCAACCCTCCACACACAGCAGAGTGGATCAACAG	352		
OY	60 ProProlalrvalThrlrrileysmetGucyAsnProAnsglnValasnGlySerArg	79		
Db	353 ---CCAGTGAAGGGTCAACGTCAACGAGGAGAT-----GACACATGAATGATCCAGG	403		
OY	80 AsnserProaspaspocysservalAlalysglylylmvetalSerSerSerAspn	99		
Db	404 GAGCTCCGGTGGACGTGACGCTTTACCAATCCACAGCTGTGGGGGAGGAGATCC	463		
OY	100 ValGlymetasnrglyserTyrmegtungluLyshis---lleProProAsmMet	118		
Db	464 AACCCCTGAACATCAACAGCTTTATGACGAGAAGAAATGGCCCCCTCTCCCAACATG	523		
OY	119 ThrThrasnglnuarghValillevalProalaasprrothlrLeutrPserThrAsphIs	138		
Db	524 ACCACCAACAGAGAGAGTCACTGCTCCCGCAGACCCCAACACTGTGGACACAGACAT	583		
OY	139 ValArgGlTrpleuenuLrrpralavalysgluytryGlyLeuProaPvalAspIleLeu	158		
Db	584 GTAGGGCAATGGGTGAAGTGGCCATTAAAGGATGACAGCTGATGAGATCGAACATCC	643		
OY	159 LeuPheglAsnalleaspglylysgluLeucySlysmecthrlyAsaPaPheglNarg	178		
Db	644 TTTTTCAGAACATGATGATGCAAGGAACTGTATAAATGAACMAAGAGACTTCCTCCG	703		
OY	179 LeutrProserTyraSnallaasplleuleuenserHisleuhstYrleuatrgLuArg	198		
Db	704 GCCACACACCTTACAAACAGAAAGTGTGTGTGCACACCTCACTGTACTCCAGGAA---	760		
OY	199 GlyAlaThrPhelleherproasnhservalYyrprogluallnrglnarglliehr	218		
Db	761 ---AGTTCAGCTGGCTATTAATCAACCTCCACACCGCAACATCTCAAGATTGAGT	817		
OY	219 ThrArpProaspLeuProtyrGlnuInlaIargtgserralatPrhrSerHisSerHis	238		

CC at fusion junctions of other observed translocations are given in
 CC AA050679-050683). NOTE: the 1890 residue sequence AA050662 does not
 CC appear in the specification; it is a combination of the Hum-Fli-1
 CC cDNA sequence (Fig 7, AA050644) and the individual intron sequences
 CC with their intron-exon junctions (Fig 13).

XX Sequence 1890 BP; 457 A; 561 C; 458 G; 414 T; 0 other;

Alignment Scores:

Pred. No.: 8,1e-118 Length: 1890
 Score: 1506.50 Matches: 310
 Percent Similarity: 62.10% Conservative: 57
 Best Local Similarity: 52.45% Mismatches: 77
 Query Match: 61.64% Indels: 147
 DB: 14 Gaps: 15

US-09-902-772-2 (1-451) x AA050662 (1-1890)

QY 5 IlelysglualaleuSerValSerGluspinSerleuPheglucysAlaTyrGly 24
 Db CTGACAGAGGCTGTGGTGTGAGCAGACAGCAGTCCCTTGAAGTCCGAGGGA 153
 QY 25 SerProHisleuAlaIySerThrGluMetThrAlaSerSerSerSerGlyTyrGly 43
 Db GCGGACGCCATCTCCCAAGCCGACATGACTCTCGGGAGTCTGACTAGGCGAG 213
 QY 44 ThrSerlyMetSerProArgValProGlnInAspThrleuSerGlnProProAlaArg 63
 Db CCCACAGACATCAACCCCTCCACACAGCAGAGAGTGCATCATCAG---CCAGTGAG 270
 QY 64 ValThrIlelyMetGluCysAsnProAsnGlnValAsnGlySer----- 78
 Db GTTACAGCTCAAGCGCGAGTAT-----GACCACATGATGATCAGGTAAGTACCCAG 324
 QY 79 -----ArgAsnSerProAsp 83
 Db GCCTGTGACAGATTGGCCTTTGGCCTTTGCCCTCTCCTTGAAGAGTCTCCGGTG 384
 QY 84 AspCysSerValAlaIyGlyGlyMetValSerSerSerSerAspAsnValGlyMetAsn 103
 Db GACTGACAGCGTTTACCAATGACAGAGTGTGTGGGAGGAGGAGTCCAAACCCATGAC 444
 QY 104 TyrIlySerTyrMetGluGluLysHis---IleProProProAsnMetThrThrAsnGlu 122
 Db TACACACAGCTATATGAGCAGAAAGATGCCCCCTCTCCCAACATGACACCAACAGAG 504
 QY 123 ArgArgValIleValPro----- 128
 Db AGGAGATCATCTGCCCGCAGGTAATTCAGAAACAGGCTGCTGGCGCCTCTTGCT 564
 QY 129 -----AlaAspProThrleuThrSerThrAspHisValArgGlnTyr 142
 Db AACACAGCTTCTCTCCAGACCCACACTGTGACACAGCATGTGAGCAATGG 624
 QY 143 LeuGluThrAlaValIyGluTyrGlyLeuProAspAlaAspIleleuPheGlnAsn 162
 Db CTGGAGGGGCCATAAAGAGATAGCTGTGATGAGATCGACACATCTTTTTCACAAAC 684
 QY 163 IleAspIlylysglueuCysLysMetThrLysAspAspPheGlnArgLeuThrProSer 182
 Db ATGGATGGCAAGAACCTGTAAATATGACAGAGAGACTTCTCCGCCACCAACCTC 744
 QY 183 TyrAsnAlaAspIleleuLeuSerHisLeuHisTyrLeuArgGlu-----ArgGly 199
 Db TACAAACAGGAAGTGTGTGTACACCTCAGTAACTCAGGAAAGTAAGTGTCCGCCA 804
 QY 200 AlaIle-----PheIlePhe----- 204
 Db AGTACCCAGGCGTGGGTATTAACCTGTATGTATTTGCTTCAGGTTCACTGCTGCC 864
 QY 205 ProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro----- 221
 Db TATATATCAACCTCCACACGACCAATCTCAGGATTTGATGTCAAGAAAGTAAGTTT 924

QY 222 -----AspLeuPro 224
 Db GTTCTTTTGTGCACTTAATAATTAATGTAACCCCTATTGTATTTGTCTATTAGACCTTCT 984
 QY 225 TyrGluGlnAlaArgSerAlaTyrThrSerHisSerHisProThrGlnSerLys--- 243
 Db TATGACTGACAGAGAGAGAGATGGGCAATTAACATGAATTTCTGGCCTCAACAAAGT 1044
 QY 244 -----AlaThrGln 246
 Db AAGTAATGTTTATAGTCTTTGAGGCTCAGCTCATCTTCTTCCCTTGTCCACAGGT 1104
 QY 247 Pro-----SerSerThrValProLysThrGluAspGlnProGln----- 261
 Db CTCTCCCTTGGAGGGGCAACAGCATGATGAAGATACAGAGCAACGCCGCCAGCCAGGT 1164
 QY 262 -----LeuAsp 263
 Db ACTGCCAGAGATATGTAATCTCTCTTTCGAAAGCAAAATTCCTTTTATTTCTCTAGAT 1224
 QY 264 ProThrGlnIleleuGlyProThrSerSerArgIleuAlaAsnPro----- 278
 Db CCGTATGATCTGTGGCCGACAGCAGATGCTGCTACACCACTGTGAGTATACCTTG 1284
 QY 279 -----GlySerGlyGlnIle 283
 Db GCCTCGAAGCTTTTGTCTCTCCGTTTCTCAGGGGTGACAGAGGAGGAGGCAATC 1344
 QY 284 GlnLeuThrGlnPheleuLeuGluLeuLeuSerAspSerSerSerAsnSerAsnCysIleThr 303
 Db CAGCTGTGGCAATCTCCCTGGAGCTCTCTCCGACAGCCCAACGCGACCTGATATCACC 1404
 QY 304 ThrGluGlyThrAsnGlyLysIlePheLysMetThrAspProAspGluValAlaIleArgTyr 323
 Db TGGAGGGGACCAAGGGGAGTTTAAATGACGAGCCCGATGAGGTGGCCAGGCGCTGG 1464
 QY 324 GlyIleArgLysSerLysProAsnMetAsnTyrAspLysLeuSerAlaIleuArgTyr 343
 Db GCGAGGGGAAACCAAGCCCAACATGATATGACAAAGTGAGCGCGCTCCCTGTAT 1524
 QY 344 TyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPheAsp 363
 Db TACTATATATAAAACATTATGACCAAGTGCAGGCAAAAGATATGCTTCAAAATTTGAC 1584
 QY 364 PheHisGlyIleAlaGlnAlaLeuGlnProHisProGlnIleuSerSerMetTyrLysTyr 383
 Db TTCACGGCATTTGCCAGGCTCTGACGACCATCCGACGAGTGTCTCATGTACAGTAC 1644
 QY 384 ProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMetAsnPheVal 403
 Db CTTCTGACATCTCTCATGATGCTCTTACCATCCACAGCAGAGAGGAACTTTGTGC 1704
 QY 404 AlaProHisProAlaIleuProValThrSerSerSerPheAlaIleProAsnPro 423
 Db CTCTCCATCATCTCCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764
 QY 424 TyrThrAsnSerProThrGlyGlyIleTyrProAsn-----ThrArgLeuProAla 440
 Db TACTGAGACCTCCCGACAGGGGAGATGTACCCCAACCCCAACGTCGCCCATCTTAAC 1824
 QY 441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451
 Db ACCACAGTGCCTTACACTAGGAGCTACTAC 1857

RESULT 7

ABA48124/c
 ID ABA48124 standard; DNA: 567-BP.

XX ABA48124;

DT 01-FEB-2002 (first entry)

DE	Human breast cell single exon nucleic acid probe #6819.
xx	
KW	Human: microarray; single exon probe; gene expression; breast;
KM	disease; cancer; ss.
xx	
OS	Homo sapiens.
xx	
PN	MO200157271-A2.
xx	
PD	09-AUG-2001.
xx	
PF	30-JAN-2001; 2001MO-US00662.
xx	
PR	04-FEB-2000; 2000US-0180312.
xx	
PR	26-MAY-2000; 2000US-0207456.
xx	
PR	30-JUN-2000; 2000US-0608408.
xx	
PR	03-AUG-2000; 2000US-0632366.
xx	
PR	21-SEP-2000; 2000US-0234687.
xx	
PR	27-SEP-2000; 2000US-0236359.
xx	
PR	04-OCT-2000; 2000GB-0024263.
xx	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
xx	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
xx	
DR	WPI: 2001-496933/54.
xx	
PT	New spatially-addressable set of single exon nucleic acid probes,
xx	
PT	useful for measuring gene expression in sample derived from human
xx	
PT	breast, comprises number of single exon nucleic acid probes
xx	
PS	Claim 4; SEQ ID NO 6819; 327bp + sequence listing; English.
xx	
CC	The invention relates to a spatially-addressable set of single exon
xx	
CC	nucleic acid probes for measuring gene expression in a sample derived
xx	
CC	from human breast and BT 474 cells. The method involves contacting
xx	
CC	the probes with a collection of detectably labelled nucleic acids
xx	
CC	derived from mRNA of human breast, and then measuring the label
xx	
CC	bound to each probe of the microarray. The probes are useful for
xx	
CC	verifying the expression of regions of genomic DNA predicted to
xx	
CC	encode proteins. They are useful for gene discovery, and for
xx	
CC	determining predisposition and/or prognosing breast disease. Gene
xx	
CC	expression analysis is useful for assessing the toxicity of chemical
xx	
CC	agents on cells. The microarray of this invention presents a far greater
xx	
CC	diversity of probes for measuring gene expression, with far less bias
xx	
CC	than expressed sequence tag microarrays. The method is suitable for
xx	
CC	rapid production of functional information from genomic sequence. The
xx	
CC	present sequence is a single exon nucleic acid probe of the invention.
xx	
CC	Note: The sequence data for this patent did not form part of the
xx	
CC	printed specification, but was obtained in electronic format directly
xx	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.
xx	
SQ	Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
xx	
Alignment Scores:	
Pred. NO.:	2,91e-71
Score:	948.00
Percent Similarity:	97.74%
Best local Similarity:	95.48%
Query Match:	38.79%
DB:	22
	Matches: 567
	Conservative: 169
	Mismatches: 4
	Indels: 0
	Gaps: 0
US-09-902-772-2 (1-451) x ABA48124 (1-567)	
OY	275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeuLeuSer 294
Db	534 CTGACCTCATTTAGGCGGTGGCCAGATCCAGCTTTGGCACTTCCTCTGAGCTCTGCG 475
OY	295 AspSerSerAsnSerAsnCysJleThrTrpGluGluGlyThrAsnGlyGluPheLeuMetThr 314
Db	474 GACAGGCTCAACATCCGCTGCTCATCACTGCGAAGGACACCAACGGGAGTTCAACATGACG 415
OY	315 AspProAspGlnValAlaArgArgTrpGlyGluArgGlySerGlyProAsnMetAsnTrp 334

Db	414	GATCCCGACGAGTGTGCCCGGGCTGGGAGAGCGGAAGACCAACCAATGAATCTAC	355
Qy	335	AspIysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis	354
Db	354	GATAGCTCAGCGCGGCCCTCCGTTACTGACTATGACAAAGACATCATGACCAAGTCCAT	295
Qy	355	GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis	374
Db	294	GGGAAGCCCTACGCCCTTACAAAGTTGCACCTCCAGGGATGCGCCAGGCCCTCCAGCCCCAC	235
Qy	375	ProProlGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis	394
Db	234	CCCCCGAGTCTATCTCTGTACAGTACCCCTAGACCTCCGTACATGGGCTTCATAC	175
Qy	395	AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer	414
Db	174	GCCCCACACAGAAAGATGAACCTTGTGGGGCCCCACCCCTCCAGCCCTCCCGTACATCT	115
Qy	415	SerSerPhePheAlaAlaProAsnProTyrTrrPaasSerProThgIyGlyIleTyrPro	434
Db	114	TCCAACTTTTTCGTCGCCCAACCCATACCTGGAATTCACCAACTGGGGGTATATACCCC	55
Qy	435	AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	451
Db	54	AACACTAGGCTCCCCACCAAGCATATGCTTCTCATCTGGGCACTTACTTAC	4
RESULT 8			
AC	ABA66003/c	ABA66003 standard; DNA; 567 BP.	
XX	ABA66003;		
XX	01-FEB-2002 (first entry)		
DE	Human foetal liver single exon nucleic acid probe #14308.		
XX	Human; foetal liver; gene expression; single exon nucleic acid probe; ss.		
XX	Homo sapiens.		
PN	WO200157277-A2.		
PD	09-AUG-2001.		
XX	30-JAN-2001; 2001WO-US00669.		
XX	04-FEB-2000; 2000US-0180312.		
PR	26-MAY-2000; 2000US-0207456.		
PR	30-JUN-2000; 2000US-0608408.		
PR	03-AUG-2000; 2000US-0633666.		
PR	21-SEP-2000; 2000US-0234687.		
PR	27-SEP-2000; 2000US-0236359.		
PR	04-OCT-2000; 2000GB-0024263.		
XX	(MOLE-) MOLECULAR DYNAMICS INC.		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
XX	WPI; 2001-483447/52.		
XX	Human genome-derived single exon nucleic acid probes useful for		
PT	analyzing gene expression in human foetal liver -		
PS	Claim 4; SEQ ID NO 14308; 639pp + sequence listing; English.		
CC	The invention relates to a single exon nucleic acid probe for		
CC	measuring human gene expression in a sample derived from human foetal		
CC	liver. The single exon nucleic acid probes may be used for predicting,		
CC	measuring and displaying gene expression in samples derived from human		
CC	fetal liver. The present sequence is a single exon nucleic acid		
CC	probe of the invention.		
CC	Note: The sequence data for this patent did not form part of the		

CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:

Pred. No.:	2,91e-71	length:	567
Score:	948.00	Matches:	169
Percent Similarity:	97.74%	Conservative:	4
Best Local Similarity:	95.48%	Mismatches:	4
Query Match:	38.79%	Indels:	0
DB:	22	Gaps:	0

US-09-902-772-2 (1-451) x ABA66003 (1-567)

QY	275	LeuAlaAspProIleSerIylInIleGlnLeuTrpGlnPheLeuLeuGlnIleuLeuSer	294
Db	534	CTGACCTCATTTAGCAGTGGCCAGATCCAGCTTTGGACGTTCTCTGGAGCTCTGTGGC	475
QY	295	AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyIleuPheLysMetThr	314
Db	474	GAGAGCTCCAACTCCAGCTGCATCCCTGGAGAGCCACCAACGGGGAGTTCCAAATGACG	415
QY	315	AspProAspGluValAlaIleArgTrpArgIylGluValGlySerLysProAsnMetAsnTyr	333
Db	414	GATCCCGACAGAGTGGCCCGCGCGCTGGGAGAGCGGAAGACCAACCATGAACTAC	355
QY	335	AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis	354
Db	354	GATTAACCTAGCGCGCCCTCCGTACTACTGTGCAAGAACATCATACCAAGATGCAT	295
QY	355	GlyLysArgTyrIleTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis	374
Db	294	GGGAAGCGCTACGCCCTACAAAGTTGCATCTCCACGGGATGCCACAGCCCTCCAGCCAC	235
QY	375	ProProGluSerSerMetLysTyrTrpProSerAspLeuProTyrMetSerSerTyrHis	394
Db	234	CCCCCGAGTCACTCTGTACAAAGTACCCTTCAGACTCCCGTACATGGGCTCCTATCAC	175
QY	395	AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer	414
Db	174	GCCCAACCCAGAAAGTGAATTTTGTGGCGCCCAACCTCCAGCCCTCCCGGTACATCT	115
QY	415	SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro	434
Db	114	TTCAGTTTTTTGTCTCCCAACCAACCATAGTGAATTCAACAACGTGGGGATATTACCCC	55
QY	435	AsnThrArgLeuProAlaAlaHisMetProSerHisIleuGlyThrTyrTyr	451
Db	54	AATACGTAGGCTCCCAACGACCATATGGCTTCTCATCTGGGCACCTTACTAC	4

RESULT 9

ABA33090/c
ID ABA33090 standard; DNA; 567 BP

AC ABA33090;

DT 23-JAN-2002 (first entry)

Probe #11556 for gene expression analysis in human heart cell sample.

KW Human; gene expression; heart; microarray; vascular system; probe
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.

05 Homo sapiens

PN WO200157274-A2

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US006666.

(MOLE-) MOLECULAR DYNAMICS INC

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488899/53

PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -

PS Claim 4; SEQ ID No 11556; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging, and
CC monitoring and proposing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from Wipo
CC at [ftp.wipo.int/pub/published_pat_sequences](http://wipo.int/pub/published_pat_sequences).

Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:

Pred. No.:	2,91e-71	length:	567
Score:	948.00	Matches:	167
Percent Similarity:	97.74%	Conservative:	4
Best Local Similarity:	95.48%	Mismatches:	4
Query Match:	38.79%	Indels:	0
DB:	22	Gaps:	0

US-09-902-772-2 (1-451) x ABA33090 (1-567)

QY	275	LeuAlaAsnProGlySerIleGlnIleuThrpglPheIleuGluIleuSer	294
Db	534	CTTACCTCATTAAGCAGTGCCAGATCCAGCTTGGCATTCCTCGAGCTTCGTGG	475
QY	295	AspSerSerAsnSerAsnCysIleThrTrpGluGluThrAsnGluGluPheIleMetThr	314
Db	474	GAGAGCTCCAACTCCACAGCTGCATACCTGGAGAGCCAAAGGGGAGTTCAAGATACG	415
QY	315	AspProAspGluValAlaIleArgTrpGlyGluArgIleuSerIleProAsnMetAsnTyr	334
Db	414	GATCCCGACAGAGGTGGCCCGCGCTGGGGAGCGCGAAGACCAACCCACATGAACATAC	355
QY	335	AspIleuSerArgIleuIleuArgTyrTyrTyrAspIleuAsnIleMetThrIleValHis	354
Db	354	GATAACTCAGCGCGCCCTCCGTTACTACTATGCAAGAATCATATGACCAAGTGCAT	295
QY	355	GlyIleuArgTyrAlaTyrIleuSerPheAspPheHisGlyIleAlaGlnAlaIleuGlnProHis	374
Db	294	GGGAACGGGTACGGCTACAGATTCGACCTTCCACGGAGATGCCGACGGCCCTCAGCCGAC	235
QY	375	ProProGluSerSerMetTyrIleTyrProSerAspIleuProTyrMetSerSerTyrHis	394
Db	234	CCCCCGAGTCACTCTGTGACAAATACCCCTCAGACCTCCGCTCATAGGGCTCTCATAC	175
QY	395	AlaHisProGluIleuMetAsnPheValAlaProHisProProAlaIleuProValIleuSer	414
Db	174	GCCACCCACAGAAATGAATTTGTGGCGCCCCACCCCTCAGGCCCTCCCGATGACATCT	115
QY	415	SerSerPhePheAlaIleProAsnProTyrTyrPAsnSerProThrGlyGlyIleTyrPro	434

```

Db      114 TCCAGTTTTTGGTCCCAACCACTGGAATTCACCACTGGGGGTATATACCC 55
QY      435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
DB      54 AACCTAGGCTCCGCCACCACTATATGCTTCATCTGAGGACTTACTAC 4

RESULT 10
AAK14426/c
ID      AAK14426 standard; DNA; 567 BP.
XX
AC      AAK14426;
XX
DT      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe SEQ ID NO: 14417.
XX
KW      Human; brain expressed exon; gene expression analysis; probe;
KW      microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW      epilepsy; cancer; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157275-A2.
XX
PD      09-AUG-2001.
XX
PE      30-JAN-2001; 2001WO-US00667.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-483446/52.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      brains -
XX
PS      Example 4; SEQ ID NO: 14417; 650bp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      brain. They can be used to measure gene expression in brain cell samples,
CC      which may enable the diagnosis and improved treatment of nervous system
CC      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC      epilepsy and cancers. The present sequence is one of the probes of the
CC      invention.
XX
SQ      Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:
Pred. No.:      2,91e-71      Length:      567
Score:          948.00      Matches:      169
Percent Similarity: 97.74%      Conservative: 4
Best Local Similarity: 95.48%      Mismatches: 4
Query Match:    38.79%      Indels:      0
DB:            22      Gaps:      0

US-09-902-772-2 (1-451) x AAK14426 (1-567)
QY      275 leuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSer 294
DB      534 CTGACCTCATTTAGGACAGGCGGACATCCAGCTTGGGCGATGCCCTGGAGACTCTCTGCG 475
QY      295 AspSerSerAsnSerAsnCysIleThrTTPcIuGlyThraSngIyGluPheLysMetThr 314

```

```

Db      474 GACAGCTCCACATCCAGCTGCATGCACCTGGGAAGGACCAACGAGGAGTTCAGATGACG 415
QY      315 AspProAspGluValAlaArgArgTyrGlyGlyAlaArgGlySerLysProAsnMetAsnTyr 334
DB      414 GATCCCGCGAGGAGGGCGCGCGCTGGGGAAGGGAAGGAAACCAACCATATGACCAAGTCCAT 355
QY      335 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 354
DB      354 GATTAAGCTCAGCCCGCCCTCCGTTACTACTATGACAAAGAACATCATATGACCAAGTCCAT 295
QY      355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
DB      294 GGAAGAGGCTACGGCTACAAAGTGGACCTCCACAGGGAGGAGCCGCTCCAGCCGAC 235
QY      375 ProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
DB      234 CCCCCGGAGTCATCTCTGTACAGTACCCCTCAACACCTCCGTACATGGGCTCTATC 175
QY      395 AlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProValThrSer 414
DB      174 GCCCACCACAGAGATGACTTGTGGCGCCCACTCCAGCCCTCCCGTACATCT 115
QY      415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 434
DB      114 TCCAGTTTTTGGTCCCAACCACTGGAATTCACCACTGGGGGTATATACCC 55
QY      435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
DB      54 AACCTAGGCTCCGCCACCACTATATGCTTCATCTGAGGACTTACTAC 4

RESULT 11
AAK40160/c
ID      AAK40160 standard; DNA; 567 BP.
XX
AC      AAK40160;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed single exon probe SEQ ID NO: 14717.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PE      30-JAN-2001; 2001WO-US00668.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488900/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human bone marrow -
XX
PS      Example 4; SEQ ID NO: 14717; 658bp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human

```

CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.

XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:

Pred. No.:	2,91e-71	Length:	567
Score:	948.00	Matches:	169
Percent Similarity:	97.74%	Conservative:	4
Best Local Similarity:	95.48%	Mismatches:	4
Query Match:	38.79%	Indels:	0
DB:	22	Gaps:	0

US-09-902-772-2 (1-451) x AAK40160 (1-567)

```

OY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294
DB 534 CTGACCTCATTAAGGAGGCGGAGATCCAGCTTGGAGTCCCTGAGCTCTGTCG 475
OY 295 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThr 314
DB 474 GACAGCTCCACTCCAGCTGCATCACCCTGGAGAGCCACCAAGGAGTTCAAGATGACG 415
OY 315 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyr 334
DB 414 GATCCCGACGAGGAGGCGCGCTGGGAGAGCGGAGCAACCAACATGATGAACTAC 355
OY 335 AspLysLeuSerArgAlaLeuArgTyrTyrTrpAspLysAsnIleMetThrLysValHis 354
DB 354 GATTAAGCTACCGCGCTCCGTTACTACTATGACAAAGACATCATATACCAAGGTCAT 295
OY 355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
DB 294 GGAAGAGCGCTACGGCTTAAGATTGACTTCACACGGAGATCGCCAGGCCCTCAGGCCAC 235
OY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
DB 234 CCCCCGGAGCATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGCGCTCATAC 175
OY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414
DB 174 GCCACCCACAGAGATGAACTTTGTGGCGCCACCCCTCCAGCCCTCCCGTGACATCT 115
OY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 434
DB 114 TCCAGTTTGTGTGGCCCCAACCACCACTGGAATTCACCACTGGGGGTATATACCC 55
OY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
DB 54 AACACTAGGCTCCGCCACGACCATATGCTTCATCTCGGCGACTTACTAC 4

RESULT 12
AAI20932/c
ID AAI20932 standard; DNA; 567 BP.
AC AAI20932;
XX
XX 12-OCT-2001 (first entry)
DE Probe #10865 for gene expression analysis in human cervical cell sample.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
XX Homo sapiens.
XX
XX MO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.

```

```

XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000US-0024263.
XX
XX (MOE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID No 10865; 487P; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SEMP). The present sequence is one such probe. The SEMPs are derived
XX from human HeLa cells. The SEMPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcf_sequences.
XX
XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2,91e-71 Length: 567
XX Score: 948.00 Matches: 169
XX Percent Similarity: 97.74% Conservative: 4
XX Best Local Similarity: 95.48% Mismatches: 4
XX Query Match: 38.79% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-902-772-2 (1-451) x AAI20932 (1-567)
XX
OY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294
DB 534 CTGACCTCATTAAGGAGGCGGAGATCCAGCTTGGAGTCCCTGAGCTCTGTCG 475
OY 295 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThr 314
DB 474 GACAGCTCCACTCCAGCTGCATCACCCTGGAGAGCCACCAAGGAGTTCAAGATGACG 415
OY 315 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyr 334
DB 414 GATCCCGACGAGGAGGCGCGCTGGGAGAGCGGAGCAACCAACATGATGAACTAC 355
OY 335 AspLysLeuSerArgAlaLeuArgTyrTyrTrpAspLysAsnIleMetThrLysValHis 354
DB 354 GATTAAGCTACCGCGCTCCGTTACTACTATGACAAAGACATCATATACCAAGGTCAT 295
OY 355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
DB 294 GGAAGAGCGCTACGGCTTAAGATTGACTTCACACGGAGATCGCCAGGCCCTCAGGCCAC 235
OY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
DB 234 CCCCCGGAGCATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGCGCTCATAC 175
OY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414
DB 174 GCCACCCACAGAGATGAACTTTGTGGCGCCACCCCTCCAGCCCTCCCGTGACATCT 115
OY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 434

```

Db 114 TCAGTTTTTGTGCTGCCCAACCATGACTGATTCACCACTGGGGGTATATACCC 55
 QY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
 Db 54 AACCTAGGCTCCGCCACGACCATATGCTTCATCTGCGGACTTACTAC 4

RESULT 13
 AA146176/c
 ID AA146176 standard; DNA; 567 BP.
 AC AA146176;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #14862 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 RM genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200157272-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID NO 14862; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

Alignment Scores:
 Pred. No.: 2,91e-71 Length: 567
 Score: 948.00 Matches: 169
 Percent Similarity: 97.74% Conservative: 4
 Best Local Similarity: 95.48% Mismatch: 4
 Query Match: 38.79% Indels: 0
 DB: 22 Gaps: 0

US-09-902-772-2 (1-451) x AA146176 (1-567)

QY 275 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 294
 Db 534 CTGACCTCATTAAGGACGAGTGCAGATCCAGCTTGGCAGTTCCTCGAGCTCCTGCG 475

QY 295 AspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyGlnPheLeuMetThr 314
 Db 474 GACGAGCTCAACCTCCATCCATCCCTGGGAAAGCACAACGGGAGTTCCAAAGATGACG 415

QY 315 AspProAspGluValAlaArgArgTrpGlyGlnArgGlySerLysProAsnMetAsnTyr 334

Db 414 GATCCGACGAGGTGGCCCGCGCTGGGAGAGGGGAAAGCAACCAACATGATGACAC 355
 QY 335 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 354
 Db 354 GATAGCTCAAGCCCGCCCTCCGTTACTACTATGACAAAGATCATGATGACCAAGGTCCAT 295

QY 355 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHis 374
 Db 294 GGAAGGCTAGCGCTCAAGATTGACTCCACGGGATGCCAGGCCCTCCACGCCCCAC 235

QY 375 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 394
 Db 234 CCCCCGAGTATCTGTGTAACATGACCCCTCAGACCTCCGCTGATGGGCTCTATAC 175

QY 395 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 414
 Db 174 GCCCACCACAGAGATGAACTTGTGGCGCCGCCACCTCCAGCCCTCCCGTGCATCT 115

QY 415 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleLeuPro 434
 Db 114 TCAGTTTTTGTGCTGCCCAACCATGACTGGAATTCACCACTGGGGGTATATACCC 55

QY 435 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
 Db 54 AACCTAGGCTCCGCCACGACCATATGCTTCATCTGCGGACTTACTAC 4

RESULT 14
 AA106643/c
 ID AA106643 standard; DNA; 567 BP.
 XX
 AC AA106643;
 XX
 DT 09-OCT-2001 (first entry)
 XX
 DE Probe #6634 used to measure gene expression in human breast sample.
 XX
 KW Probe; human; breast disease; breast cancer; development disorder; ss;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX
 OS Homo sapiens.
 XX
 PN W0200157270-A2.
 PD 09-AUG-2001.
 XX
 PF 29-JAN-2001; 2001WO-US00661.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-476286/51.
 XX
 PT Novel single exon nucleic acid probe used to measuring gene expression
 PT in a human breast -
 XX
 PS Claim 25; SEQ ID NO 6634; 322bp; English.
 XX
 CC The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:18:59 ; Search time 60.6835 Seconds
(without alignments)
2279.222 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444

Sequence: 1 MASTKEALSVSEDSLFE.....IYPNRLPAHMPHSHGTTY 451

Scoring table:

BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09902772/rmlna.23072003_093657_14904/app_query.fasta.1.1294
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIC=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09902772.cgn2_1.1_55_etunal.23072003_093657_14904 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444	100.0	1447	4	US-08-878-177-1
2	2414.5	98.8	1528	4	US-08-878-177-3
3	1624.5	66.5	2938	2	US-08-343-443B-3
4	511	20.9	1752	4	US-09-360-779-1
5	511	20.9	1752	4	US-09-435-335-1
6	449	18.4	1604	1	US-08-306-691B-43
7	449	18.4	1604	5	PCT-US93-06251-9
8	445.5	18.2	2268	2	US-09-344-579-1
9	342.5	14.0	2667	2	US-08-469-412A-1
10	342.5	14.0	2667	2	US-09-021-715-1
11	322	13.2	2064	3	US-08-875-944B-1
12	322	13.2	2064	4	US-09-116-049-3

13	306	12.5	2410	2	US-08-780-835B-1	Sequence 1, Appl1
14	306	12.5	2410	4	US-09-303-268-1	Sequence 1, Appl1
15	306	12.5	2410	4	US-09-116-049-1	Sequence 1, Appl1
16	301.5	12.3	328	2	US-08-343-443B-5	Sequence 5, Appl1
17	301	12.3	1905	2	US-09-055-113-2	Sequence 2, Appl1
18	291.5	11.9	2266	2	US-09-213-767-1	Sequence 1, Appl1
19	273.5	11.2	5510	3	US-09-009-913-3	Sequence 3, Appl1
20	261.5	10.7	5427	3	US-09-009-913-2	Sequence 2, Appl1
21	261.5	10.7	5667	3	US-09-009-913-4	Sequence 4, Appl1
22	254.5	9.6	2544	2	US-08-469-412A-6	Sequence 6, Appl1
23	234.5	9.6	2544	4	US-09-021-715-6	Sequence 6, Appl1
24	231.5	9.5	848	3	US-09-009-913-338	Sequence 338, App
25	231.5	9.5	2280	3	US-09-009-913-8	Sequence 8, Appl1
26	231.5	9.5	2428	3	US-09-009-913-6	Sequence 6, Appl1
27	231.5	9.5	2498	3	US-09-009-913-10	Sequence 10, Appl1
28	229	9.4	852	4	US-09-020-956-44	Sequence 44, Appl1
29	229	9.4	852	4	US-09-030-607-44	Sequence 44, Appl1
30	229	9.4	852	4	US-09-605-785-44	Sequence 44, Appl1
31	229	9.4	852	4	US-09-439-313-44	Sequence 44, Appl1
32	229	9.4	852	4	US-09-352-616A-44	Sequence 44, Appl1
33	229	9.4	852	4	US-08-232-149A-44	Sequence 44, Appl1
34	227.5	9.3	1920	1	US-08-746-789A-1	Sequence 1, Appl1
35	211.5	8.7	2975	1	US-08-368-281-1	Sequence 1, Appl1
36	211.5	8.7	3240	1	US-08-368-281-3	Sequence 3, Appl1
37	181	7.4	65042	4	US-09-784-316-3	Sequence 50, Appl1
38	166	6.8	1364	5	PCT-US93-06251-65	Sequence 65, Appl1
39	166	6.8	50937	4	US-09-428-517-1	Sequence 1, Appl1
40	118.5	4.8	5194	2	US-09-599-652-1	Sequence 1, Appl1
41	117	4.8	5194	4	US-08-642-846-1	Sequence 1, Appl1
42	117	4.8	5194	2	US-09-264-604-1	Sequence 1, Appl1
43	115.5	4.7	2626	1	US-08-156-020-5	Sequence 5, Appl1
44	115.5	4.7	2626	1	US-08-156-020-3	Sequence 3, Appl1
45	114.5	4.7	2626	1	US-08-156-020-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-878-177-1
Sequence 1, Application US/08878177
Patent No. 6294354
GENERAL INFORMATION:
APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
TITLE OF INVENTION: the Proteins
FILE REFERENCE: Chugai Selyaku Kabushiki Kaisha 5001
CURRENT FILING DATE: 1997-06-18
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1
LENGTH: 1447
TYPE: DNA
ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
US-08-878-177-1

Alignment Scores:

Pred. No.: 4.05e-247 Length: 1447
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-902-772-2 (1-451) x US-08-878-177-1 (1-1447)

QY 1 MetAlaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 63 ATGGCAAGCAGCTATTAAAGAGCACTTATCACTGGTGAAGCAAGCAGCTCTTTTGAG 122
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB 123 TGTGCTTACGGATGCGCCCACTTGCAGAAAGCAAAATGACAGCCCTCTTCACAGAA 182

QY 41 TyrGlnThrSerLysMetSerProAlaValProGlnInaSprLysSerGlnPro 60
 Db 183 TATGGCAAAATCAAAAGATGAGCCCGCGTTCCCGACAGAGACTGTTATCAACAGCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 243 CCGGCCAGAGTTACCATTAAGATGAGTGTACCCAAACCAAGTTAATGGGTCAAGGAAP 302
 QY 81 SerProAspCysSerValAlaLysGlyLysMetValSerSerSerAspVal 100
 Db 303 TCACCTGATGATGACGACCGTGGCAAAAGAGGAAATGGTTAGCACTTACAGCAATGTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluLysHisIleProProProAsnMetThrThr 120
 Db 363 GGGATGAACATATGGAAGCTCATGGAAGACATATTCGGCTCCAAATATGACAAACC 422
 QY 121 AsnGluArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
 Db 423 AATGAACGAAGATTTATGTGCCAGCATCTTACCTTATGAGCAGACACCATGTAAGG 482
 QY 141 GlnTyrLeuGlnTyrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 Db 483 CAGTGGCTGAGTGGGAGTGAAGAGATGATGCTTCCAGACGTCATCTTGTGTTTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 543 CAGAACATTCATGGGAAAGAGTGTGTAATGACCAAAAGATGACTTCCAGAGACTCAGC 602
 QY 181 ProSerTyrAsnIleAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgValAla 200
 Db 603 CCGAGCTATACCAATATCTCTGCTGACACTTACACTTACCTGAGAGAGAGAGAGGCC 662
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArg 220
 Db 663 ACTTTATTTTCCAAATACATCATGTTTACCCAGAAAGCAAGCAAGAAATACAAACAGG 722
 QY 221 ProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrpThrSerHisSerHisProThr 240
 Db 723 CCGAGTTTACCTTATAGCAAGAGAGAGATCAGCTGCGAGTCCAGACCCCTCCACT 782
 QY 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlnAspGlnArgPro 260
 Db 783 CAGTCAAAAGCTTACCAACATCATCTTCAACAGTCCCAAAACAGAACGCCAGCTCT 842
 QY 261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
 Db 843 CAGTTAGATCCTTATCAGATTCTTGGACCGACAGACGCGCTTCAATCCAGGAGT 902
 QY 281 GlyGlnIleGlnLeuThrPheGlnPheLeuGlnLeuLeuSerAspSerSerAsnSerAsn 300
 Db 903 GGGCAGATACAGCTATAGGAGTTCCTACGTGAGACTTCTCGGACAGCTCCCAACTCCAC 962
 QY 301 CysIleThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320
 Db 963 TGCATACCTGGAGAGGACAAATGGGAGTTCAAATGACAGACACCTGTATGAGTGGCT 1022
 QY 321 ArgArgTyrGlyGlnArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340
 Db 1023 CCGCGTTGGGAGAGAGAAAGCAAACTACATGACATGACAACTCAGCGCTGCA 1082
 QY 341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360
 Db 1083 CTTCGCTACTACTATACAAAAATATATGACTAAAGTTCATGCTAAAGCTATGCTCTAC 1142
 QY 361 LysPheAspPheHisGlyLysIleAlaGlnAlaLeuGlnProHisProProLysSerSerMet 380
 Db 1143 AATTTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCCTCCACAAATATCATCTCAGT 1202
 QY 381 TyrLysTyrProSerAspLeuProTyrMetLysSerTyrHisAlaHisProGlnLysMet 400
 Db 1203 TACAATATACCATGAGACCTCCCTACATGAGTTCCTACATGACACACCCCAAGAGATG 1262

QY 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAla 420
 Db 1263 AACTTTAGCTCCCTCCATCCCTGCTTGGCCGTACCTCATCCACTTTTGTGTCG 1322
 QY 421 ProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuProAla 440
 Db 1323 CCTAATCCATATGGAATTCACCAACTGGAGGATCTACCCCAATACCAAGCTGCGAGCT 1382
 QY 441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451
 Db 1383 GCATCATATGCTTCCCATCTTGGCAGCTACTAC 1415
 RESULT 2
 US-08-878-177-3
 ; Sequence 3, Application US/08878177
 ; Patent No. 6294354
 ; GENERAL INFORMATION:
 ; APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwanoto et
 ; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
 ; FILE REFERENCE: Chugai selyaku kabushiki kaisha 5001
 ; CURRENT APPLICATION NUMBER: US/08/878,177
 ; CURRENT FILING DATE: 1997-06-18
 ; NUMBER OF SEQ. ID NOS.: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ. ID NO. 3
 ; LENGTH: 1528
 ; TYPE: DNA
 ; ORGANISM: c-ery gene, chicken DNA
 ; US-08-878-177-3
 Alignment Scores:
 Pred. No.: 5,57e-244 Length: 1528
 Score: 2414.50 Matches: 450
 Percent Similarity: 94.14% Conservative: 0
 Best Local Similarity: 94.14% Mismatches: 1
 Query Match: 98.79% Indels: 27
 DB: 4 Gaps: 1
 US-09-902-772-2 (1-451) x US-08-878-177-3 (1-1528)
 QY 1 MetAlaSerThrIleLysGlnAlaLeuSerValValSerGlnAspGlnSerLeuPheGln 20
 Db 63 ATGGCAAGCACTATTAAGAGAGATGATGAGTGTAGTGAAGACCAAGTCTTGTGAG 122
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGln 40
 Db 123 TGGGCTAGGAGTGGCCGACCTTGAAGAGCAAGAAATGACAGCTTCTTCCATGAA 182
 QY 41 TyrGlnThrSerLysMetSerProAlaValProGlnInaSprLysSerGlnPro 60
 Db 183 TATGGCAAAATCAAAAGATGAGCCCGCGTTCCCGACAGAGACTGTTATCAACAGCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 243 CCGGCCAGAGTTACCATTAAGATGAGTGTACCCAAACCAAGTTAATGGGTCAAGGAAP 302
 QY 81 SerProAspCysSerValAlaLysGlyLysMetValSerSerSerAspVal 100
 Db 303 TCACCTGATGATGACGACCGTGGCAAAAGAGGAAATGGTTAGCACTTACAGCAATGTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluLysHisIleProProProAsnMetThrThr 120
 Db 363 GGGATGAACATATGGAAGCTCATGGAAGACATATTCGGCTCCAAATATGACAAACC 422
 QY 121 AsnGluArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
 Db 423 AATGAACGAAGATTTATGTGCCAGCATCTTACCTTATGAGCAGACACCATGTAAGG 482
 QY 141 GlnTyrLeuGlnTyrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 Db 483 CAGTGGCTGAGTGGGAGTGAAGAGATGATGCTTCCAGACGTCATCTTGTGTTTC 542

```

QY 161 GlnAsnIleAspGlyLeuGlyLeuGlySerHisLeuHisLeuGlyArgLeuThr 180
Db 543 CAGAACATTGATGGAGAAAGTTGTGTAATAATGACCAAGATGCTCCAGAGACTCAGC 602
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisLeuGlyArgLeuGlyArg 198
Db 603 CCGAGCTATTAACGAGATATCTCTCTACACCTACACTACACTACAGAGAGACTCTCT 662
QY 198 ----- 198
Db 663 CCACATTGACTTCAGATGATGTTGATAGAGCCCTTACAAAACCTCCACGGTTATGATCAT 722
QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213
Db 723 GCTAGAAACACAGAGAGAGCCACTTTATTTTCCAAATACATCAGTTTACCCGAAAGCA 782
QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233
Db 783 ACGCAAGAAATTAACACAGAGCCAGATTTTACCTTATGAGCAAGAGAGATCAGCGTGG 842
QY 234 ThrSerHisSerHisProThrGlnSerTyrAlaThrGlnProSerSerSerThrValPro 253
Db 843 ACGAGTACAGCCATCCACTCAGTCAAAAGCTACCAACCATCATCTTCAACAGTGCC 902
QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
Db 903 AAAACAGAACAGACCGCTCTCCTCAGTTGATCCTTATCAGATCTTGAGACCGACGACGAC 962
QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeuLeu 293
Db 963 CGCTTTCGAAATCCAGGAGAGTGGGACATACAGTATGAGCAGTTCCTACTGAGACTCTCTG 1022
QY 294 SerAspSerSerAsnSerAsnGlyIleThrTyrGlnIleThrAsnGlyIlePheLeuMet 313
Db 1023 TCGGACAGCTCCAACTCACTGATCACCCTGGAGGACCAATGGGGAGTTTCAGATG 1082
QY 314 ThrAspProAspGluValAlaArgArgTyrGlyIleGlySerSerSerProAsnMetAsn 333
Db 1083 ACAGACCTCATGAAAGGCTCGGCGCTTGGGGAGAGAGGAGAAAGCAAACTTACATTAAC 1142
QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
Db 1143 TATGACAACTCAGACCGCTGACCTCGCTACTACTATGACAAATAATATTATGACTAAGTT 1202
QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisIleAlaGlnAlaLeuGlnPro 373
Db 1203 CATGGTAAACGCTATGCTCAAAATTTGATTTCACGGAATCGCTCAGGCCCTCCACCT 1262
QY 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
Db 1263 CACCCCTCCAGATCATCATCATGATCAATACCATCAGACCTCCCTCATGATGATTCCTTAC 1322
QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProValThr 413
Db 1323 CATGCAACACCCCAAGATGAACTTGTAGCTCCCATCCCTCGCTTGGCCGCTAAC 1382
QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrPheSerProThrGlyIleTyr 433
Db 1383 TCATCCAGCTTTTGTGCTGCCCTTAATCATCATGGAATTCACCACTGGAGGATTCAT 1442
QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1443 CCCAATTCAGAGCTGCCAGCTCATATGCTTCCATCTTGGACCTTACTTAC 1496

```

```

APPLICANT: Peter, Martine
APPLICANT: Ploongastel, Beatrice
APPLICANT: Thomas, Gilles
APPLICANT: Zucman, Jessica
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
TITLE OF INVENTION: TRANSLOCATIONS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
City: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989,6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2938 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 143..1498
US-08-343-443B-3
Alignment Scores:
Pred. No.: 1,696-160 Length: 2938
Score: 1624.50 Matches: 306
Percent Similarity: 79.43% Conservative: 57
Best Local Similarity: 66.96% Mismatches: 83
Query Match: 66.47% Indels: 11
Gaps: 7
US-09-902-772-2 (1-451) x US-08-343-443B-3 (1-2938)
QY 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 143 ATGGACGGGACTATTTAAGAGGCTCTGTGCTGGTGGACAGCAGCCAGTCCCTTTGAC 202
QY 21 CysAlaTyrGly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
Db 203 TCAGCGTAGCGAGGAGGAGCCCATCTCCCAAGGCCAGCAATGCTGCTGGGAGTCT 262
QY 40 GluTyrGlyLinhThrSerLysMetSerProArgValProGlnGlnAspTyrPheLeuSerGln 59
Db 263 GACTACGGGAGGCCCAAGAGATCAACCCCTCCACACGACGAGGAGTGCATCAATCAG 322
QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79

```

```

Db      323 ----CCAGTGAAGGCTAACGTCACAGGAGTAT-----GACCCACATGAATGATCCACGG 373
Qy      80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db      374 GAGTCTCGGGGAGCTGAGCGGTATGACAAATGCAGCAAGCTGGGGCGAGGGAGTCC 433
Qy      100 ValGlyMetAsnTyrGlySerTyrMetGluGlyLysHis---LleProProAsnMet 118
Db      434 AACCCCATGAATCTCAACAGCTATATGACAGAAAGATGGCCCCCTCCCTCCACATG 493
Qy      119 ThrThrAsnGluArgValIleValProAlaAspProThrLysLeuTyrSerThrAspHis 138
Db      494 ACCACCAACGAGAGAGATGATGTCTCCGCCACACCCACACCTGTGTGACACAGAGAT 553
Qy      139 ValArgGlnTyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeu 158
Db      554 GTGAGCGAATGGCTGGAGTGGGCTAAAGAGATGATGCTTGAAGAGATGACACATCC 613
Qy      159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
Db      614 TTTTCCAGACATGATGATGCAAGGAAGCTGTAAATGAACAAGAGAGACTTCTCCGC 673
Qy      179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLysHisTyrLeuArgLysArg 198
Db      674 GCCACCAACCTCTCAACACAGAGAGTGTGTGTACACACTGATACCTACCTACAGGAA--- 730
Qy      199 GluValaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr 218
Db      731 ---AGTTCACTGCTGGCTTAATATACAACTCCACACCCGACACATCTCCACGATGGT 787
Qy      219 ThrArgProAspLeuProTyrGluGlnAlaArgArgSerLysIleThrSerHisSerHis 238
Db      788 GTCAAGAGAGACCTTGTATGACTGACAGAGAGAGAGCTGGGCAATGACATGAT 847
Qy      239 ProThrGlnSerLysAlaThrGln---ProSerSerSerThrValProLysThrGluAsp 257
Db      848 TCTGGCTCAACAAAGTCTCCCTCGGAGGGGACCAAAAGATGATGATGATGATGATGAT 907
Qy      258 GluArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
Db      908 CAACGGCCCGCAGGATGCTGATACATGATCTGGGCCCGACACACAGTCCGCTAGCCAAAC 967
Qy      278 ProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerSerSerSer 297
Db      968 CTGGAAGCGGCGGATGCAAGCTGTGCAATCTCTCGGAGCTGCTCCGACAGCGCC 1027
Qy      298 AsnSerAsnCysIleThrTyrGluGlyThrAsnGlyLysPheLysMetThrAspProAsp 317
Db      1028 AACGCCAGCTGATCACCTGGAGGAGGACCAAGGGGATTCAAATGAGCGAGCCCGAT 1087
Qy      318 GluValaAlaArgArgTyrGlyLysArgLysSerLysProAsnMetAsnTyrAspLysLeu 337
Db      1088 GAGGTGGCGAGCGCTGGGCGAGCAAGCAAGCCAAACATGAAATTCAGCAAACTG 1147
Qy      338 SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArg 357
Db      1148 AGCGGGGCTCCCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
Qy      358 TyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisSerProGlu 377
Db      1208 TATGCTTACAAATTTGACTTCCACAGGATGGCCAGGCTGTGACCCACATCCAGCAG 1267
Qy      378 SerSerMetLysTyrTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisPro 397
Db      1268 TCGTCATGTACAGTACCTTGCATCATCTCCATCATGCTCCATCATGCTCCACACAG 1327
Qy      398 GluLysMetAsnPheValAlaProHisSerProAlaLeuProValThrSerSerSerPhe 417
Db      1328 CAGAGGCTGAACTTGTCTCCCTCCCATTCATCTCCATGCTGCTGCTCCATGCTCCAGCTTC 1387
Qy      418 PheAlaAlaProAsnProTyrTyrPheSerProThrGlyLysIleTyrProAsn----- 435

```

```

Db      1388 TTGGAGCCCGCATACATACTGAGACCTCCCCACAGGGGGGATCTACCCACCCAC 1447
Qy      436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      1448 GTCCCCCGCATCTTAACACCCACAGTGTGCTTACACTTACAGCAGCTACTAC 1498

RESULT 4
US-09-360-779-1
; Sequence 1, Application US/09360779
; Patent No. 6268216
; GENERAL INFORMATION:
; APPLICANT: Denaris, Evan S.
; APPLICANT: Evodaro, Dmitry V.
; APPLICANT: Hendricks, Timothy J.
; TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
; FILE REFERENCE: CASE-03828
; CURRENT APPLICATION NUMBER: US/09/360,779
; EARLIER FILING DATE: 1999-07-26
; EARLIER APPLICATION NUMBER: 60/094,264
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1131)
US-09-360-779-1

Alignment Scores:
Pred. No.: 7.4e-44 Length: 1752
Score: 511.00 Matches: 123
Percent Similarity: 50.69% Conservative: 23
Best Local Similarity: 42.71% Mismatches: 63
Query Match: 20.91% Indels: 79
Gaps: 10

US-09-902-772-2 (1-451) x US-09-360-779-1 (1-1752)
Qy      209 ValTyrProGluAlaThrGlnArgIleThrThrProAspLeuProTyrGluGlnAla 228
Db      369 GTGCCCCCGCTGGGCGTCCCGCCCATCTCCACCGCCA----- 407
Qy      229 ArgArgSerAlaTyrThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSer 248
Db      408 -----GTCCCCGCGCATGACAGACAGAGCGGACGCTCCAGCCGTGCT 452
Qy      249 SerSerThrValProLysThrGluAspGlnArg-----Pro 260
Db      453 GATCACATGATCTTACACATGCTCCGTGGAGATGCTTTTAAAGAAAGGAAGACCC 512
Qy      261 GlnLeu-AspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
Db      513 GAGCTGGGGGCGC-----CTGAGCCCTCGCGGTACAGAA-----GGCAG 551
Qy      280 rGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAs 300
Db      552 CGGCGACATGCAGATTGTGGCAGATTTCATGAGACTCTGTGACACACCGCGGACGCCGG 611
Qy      300 nCysIleThrTyrGluGlyThrAsnGlyLysPheLysMetThrAspProAspGluValAla 320
Db      612 CTGATGTGCGGGGAGGGCGCGCAGCGGAGTTTCACCTGACCCACCGCGAGGCTGGC 671
Qy      320 aArgArgTyrGlyLysArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAl 340
Db      672 GCGACGCTGGGGCGAGCGCAAGACCAAGCCCAATATGATACACAGCTAACTCAGC 731
Qy      340 aLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360
Db      732 ACTGCGCTACTACTAGACAAACAAACATGATGAGCAAGGTGACGCGCAAGGCTACGCTA 791

```

```

QY 360 rlyspheasphhisgllylealaglalaLeuGlnProHisPro----- 375
Db 792 CGGCTTGACTTCCAGGAGCCGCGACAGCTTGCCACCCGCGACGCCACGC 851
QY 376 -----ProGlnSerSerMetYrlySyrPr 384
Db 852 CGCGCGTCCGCGCGCGCGAGCGCGCGCGCGAGATGCGCAGCTTTCAGAGTCC 911
QY 384 oSerAspLeu-----ProYrMetSerSerYrHisAlaHisProGlnly 399
Db 912 GCGGCTGCTGCTCCACTGCGCTTCCCGCGCTCTCC-----AA 950
QY 399 smetAspHeValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAl 419
Db 951 ACTACACCTTATGACAGCTCGCGCGCGCGCGCGCGCTGCTCT----- 999
QY 419 alaProAsnProYrTrpAsnSerPro-----ThrGlyG1 431
Db 1000 -----TACTGGCGCTGGTCCCAACGCGACGCGCGCGCGCGCGCGCTGC 1046
QY 431 ylleYrProAsnThrArgLeuPro-----AlaAlaHisMetPr 444
Db 1047 GCTCTACCCAAACCCCGGCGCTTGCGAGCCCTCCCGCGCGCTTGGCGGCGGTGCGCGCGC 1106
QY 444 oSerHisLeuGlyThrYrYr 451
Db 1107 TTCGCACTTGGGGGCTCATTTAT 1128

```

RESULT 5

```

US-09-435-335-1
: Sequence 1, Application US/09435335
: Patent No. 6384204
: GENERAL INFORMATION:
: APPLICANT: Demeris, Evan S.
: APPLICANT: Eyodoro, Dmitry V.
: APPLICANT: Hendricks, Timothy J.
: TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
: TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
: FILE REFERENCE: CASE-04027
: CURRENT APPLICATION NUMBER: US/09/435,335
: EARLIER FILING DATE: 1999-11-05
: EARLIER APPLICATION NUMBER: 09/360,779
: EARLIER FILING DATE: 1999-07-26
: NUMBER OF SEQ. ID NOS: 23
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)..(1131)
: US-09-435-335-1

```

Alignment Scores:

```

Pred. No.: 7.4e-44 Length: 1752
Score: 511.00 Matches: 123
Percent Similarity: 50.69% Conservative: 23
Best Local Similarity: 42.71% Mismatches: 63
Query Match: 20.91% Indels: 79
DB: 4 Gaps: 10

```

```

US-09-902-772-2 (1-451) x US-09-435-335-1 (1-1752)
QY 209 ValTyProGlnAlaThrGlnArgIleThrThrArgProAspLeuProYrGlnGlnAla 228
Db 369 GTGGCCCCCGGTGGGCGTCCCGCCATCTCCACCGGCCA----- 407
QY 229 ArgArgSerAlaTrpTrpSerHisSerHisProHisGlnSerIysAlaThrGlnProSer 248
Db 408 -----GTCCCGCGGAGCGAGAGACAGAGCGGACGCTCCAGCCCGCTGCT 452

```

```

QY 249 SerSerThrValProIysThrGlnAspGlnArg-----Pro 260
Db 453 GATCAACATGTACTTACAGATCCCGCTGGAGATGCTTTTAAAGAGGAGAGAGCC 512
QY 261 GlnLeu-AspProYrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySe 280
Db 513 GAGCTGGGGGCGG-----CTGAGCGCTCGGGGTACAGAA-----GGCAG 551
QY 280 rGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAs 300
Db 552 CGGCGAGATCCAGTTGTGGCGAGTTTCTACTGAGCTGTGGCAGACCGCGGAACGCGCG 611
QY 300 nGySIIleThrPrpGlnGlyThrAsnGlyLupPheLysMetThrAspProAspGluValAl 320
Db 612 CTGCATGCGGTGGAGGAGGCGCGCCACGCGAGTTCAAGCTCACCAACCCCGAGAGGTGCG 671
QY 320 AARGATrPGLyGlnArgIysSerIysProAsnMetAsnYrAspIysLeuSerArgAl 340
Db 672 GCGAGCGTGGGGCGAGCGCAAGAGCAAGCCCATATGACTACAGACAAAGCTAAGTCGAGC 731
QY 340 AleuArgIyTrYrYrAspIysAsnIleMetThrLysValHisGlySArgYrAlaTy 360
Db 732 ACTCGCTACTACTACAGCAAAACATCATGAGCAAGGTGACGCGAAGCGCTACGCTTA 791
QY 360 rlyspheasphhisgllylealaglalaLeuGlnProHisPro----- 375
Db 792 CGGCTTGACTTCCAGGAGCCGCGACAGGCTTGCCAGCACCCGCGCGCGCGCCACGC 851
QY 376 -----ProGlnSerSerMetYrlySyrPr 384
Db 852 CGCGCGTCCGCGCGCGCGCGAGCGCGCGCGCGCGAGATGCGCAGCTTTCAGAGTCC 911
QY 384 oSerAspLeu-----ProYrMetSerSerYrHisAlaHisProGlnly 399
Db 912 GCGTGTCTGTGCTCCACTGCGCTTCCCGCGCTCTCC-----AA 950
QY 399 smetAspHeValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAl 419
Db 951 ACTACACCTTATGACAGCTCGCGCGCGCGCGCGCGCTGCTCT----- 999
QY 419 alaProAsnProYrTrpAsnSerPro-----ThrGlyG1 431
Db 1000 -----TACTGGCGCTGGTCCCAACGCGACGCGCGCGCGCGCGCGCTGC 1046
QY 431 ylleYrProAsnThrArgLeuPro-----AlaAlaHisMetPr 444
Db 1047 GCTCTACCCAAACCCCGGCGCTTGCGAGCCCTCCCGCGCGCTTGGCGGCGGTGCGCGCGC 1106
QY 444 oSerHisLeuGlyThrYrYr 451
Db 1107 TTCGCACTTGGGGGCTCATTTAT 1128

```

RESULT 6

```

US-08-306-691B-43
: Sequence 43, Application US/08306691B
: Patent No. 5734039
: GENERAL INFORMATION:
: APPLICANT: Calabretta, Bruno
: APPLICANT: Skorski, Tomasz
: TITLE OF INVENTION: ANTISENSE
: TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seidel, Gonda, Lavoraga & Monaco, P.C.
: STREET: Two Penn Center, Suite 1800
: CITY: Philadelphia
: STATE: Pennsylvania
: COUNTRY: U.S.A.
: ZIP: 19102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: MS-DOS

```

```

SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1604 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-306-691B-43

Alignment Scores:
Pred. No.: 2,1e-37 Length: 1604
Score: 449.00 Matches: 114
Percent Similarity: 43.88% Conservative: 51
Best Local Similarity: 30.32% Mismatches: 89
Query Match: 18.37% Indels: 122
Gaps: 9

US-09-902-772-2 (1-451) x US-08-306-691B-43 (1-1604)
QY 119 ThrThrasnsluargvalilevalproalasprrhrlleutrpserthrphs 138
Db 447 ACTAAGAGACAGACACTGGGGATCCCAAGAGCCCGGACGTGAGAGAAACCAT 506
QY 139 Valarglntrpreuglutrpalavalluylglutrglyleuproaspvalaspilleu 158
Db 507 GTTCGGGACGTGGGTGATGGGCTGTGAATGATTCAGCCTGAAGGTGTAGACTTCAG 566
QY 159 Leuphglasnilleaspilysgluleucysleuthrlyspaspspneclnarg 178
Db 567 AAGTTC---TGATGATGAGAGAGAGCCCTCGCGCCCTGGGTAAAGACTCTTCTCGAG 623
QY 179 Leuthrproserlyrinalaspilleuleuserthisleuthrlysluarglu--- 197
Db 624 CTGGCCCCAGACTTGTGGGACATCTTATGGAGACATCTAGAGATCCGCAAGAGAG 683
QY 198 -----ArglyAlaThrPheIlePheProAsnThr----- 207
Db 684 GATGTGAACCATATCAAGTTAATGAGTCAACCCAGCCATATCAGAAATCCCGTATACC 743
QY 207 ----- 207
Db 744 TCGGATTACTTCACTAGCTATGATAGCATGCCAGTGTGTCACCATCGAGATTC 803
QY 208 -----SerValyProgluAlaThrGlnArg 216
Db 804 TCAGAGCCAGCTTCATCAGAGAGTCTTATCAGACGCTCATCCATCAGCTCGGAGAG 863
QY 217 Ile----- 217
Db 864 CTCTCTCCCTCAAGTATGAGAAATGATACCCCTGGTCAATTCTCGAGACCCCTCTCAG 923
QY 217 ----- 217
Db 924 ACAGACACTTGCAGATGACTTGTATCAACAAGAGAGTGTGACCCAGAGACAC 983
QY 218 -----ThrThrargProaspLeuProtyrLglnAlaArgArgSer--- 231
Db 984 ATGTGATGGGAGAGACAGTGTGTAAACTCGGGGGGAGAGACTCTTTTGAAGCATATA 1043

```

```

QY 232 ---AlatrphrserhisserhisprothrGlnserlysalthrGlnProSerSer 250
Db 1044 GAGAGCTACATGATGTGTATGCCCTCACCAGAGCTCGAGAGCAGCATCATCTTTCAC 1103
QY 251 ThrValProLys-----ThrLysp----- 257
Db 1104 AGCCTGAGGCTGTCTCCCTCATATGACAGCTTCGACATGAGAGCATATCCGCTCC 1163
QY 258 -----GlnArgProGlnLeuAsp 263
Db 1164 CCACACCAAGCCCAAGGACCTTCACAGCATATGTGGGAGCCGTGTGACTTCAT 1223
QY 264 ProtyrGlnleuLeuGlyProthrSerSerArgLeuAlaAsnProGlySerGlyGlnIle 283
Db 1224 AAGGACAGGCTGTATCTGTGTGCTGCCCTAGCTGCTGATACACAGGACATGACCAATC 1283
QY 284 GlnleuThrGlnPheleuLeuGlnleuLeuSerSerSerSerSerSerSerSerSer 303
Db 1284 CAGCTATGCGAGTTCCTTCTGTGAATTAATCAGTATTAATCCTGTAGTCTTTTATCAGC 1343
QY 304 TyrGlnGlyThrAsnGlyLuphlyMetThrAspProaspGluValAlaArgArgTrp 323
Db 1344 TGACACAGAGATGGCTGGGATTCAACTTGTGACCCAGATGAGGTGGCCAGAGATGG 1403
QY 324 GlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyr 343
Db 1404 GGAAGAGGAAAAAACAACCTAGATGATTAATGAGAAATGAGCGGTGCGCTACAGCTAC 1463
QY 344 TyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPhe--- 362
Db 1464 TATTGACACAAAAACATCATCCACACAGCGGGGAAAGCATGATGATGATGATGATG 1523
QY 363 ---AspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
Db 1524 TGTGACTG-----CAGAGCTCTGTGGGGTACACCCCTGAG 1559

RESULT 7
PCT-US93-06251-9
Sequence 9, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1604 base pairs
TYPE: nucleic acid

```



```

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-9

Alignment Scores:
Pred. No.: 2.1e-37 Length: 1604
Score: 449.00 Matches: 114
Percent Similarity: 43.88% Conservative: 51
Best Local Similarity: 30.32% Mismatches: 89
Query Match: 18.37% Indels: 122
DB: 5 Gaps: 9

US-09-902-772-2 (1-451) x PCT-US93-06251-9 (1-1604)
QY 119 ThThrAsnGlnArgValIleValIProAlaAspProThrLeuTrpSerThrAspHis 138
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 447 ACTAAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 506

QY 139 ValArgGlnTrpLeuGlnTrpAlaValIleGlnLeuGlnLeuProAspValAlaPheLeu 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 GTTCGGGAGCTGGGTGATGGGTGGGTGATGATGATGATGATGATGATGATGATGATGAT 566

QY 159 LeuPheGlnAsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArg 178
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 AGATTC---TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623

QY 179 LeuThrProSerLysAsnAlaAspIleLeuLeuSerHisLeuHisLysIleValArgGlu 197
   ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 CTGGCCCAAGACATTTGTTGGGACATCTTATGGAACTTAAGATCCCTGACAGAAAGAG 683

QY 198 -----ArgGlyAlaThrPheIlePheProAsnThr----- 207
   ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 684 GATGTGAACCATATCAATGATTAATGAGTCAACCCAGCTATTCAGATCCCGCTATACC 743

QY 207 ----- 207
   ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 744 TCGGATTACTTCAATTAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 803

QY 208 -----SerValIleProGlnAlaThrGlnArg 216
   ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 804 TCAGAGCCAGCTTCATCAGACAGTCTATCAGACGCTCATCCATCCAGCTGGAGAGAG 863

QY 217 Ile----- 217
   ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 864 CTCTCTCCCTCAAGTATGAGAAATGACTACCCCTCGTCACTTCGAGACCCCTTCACAG 923

QY 217 ----- 217
   ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 924 ACAGACACCTTGAGAGATGACTTTCATCAACAAGACAGTCCGACCCAGACAC 983

QY 218 -----ThrThrArgProAspLeuProGlyGlnAlaIleArgSer--- 231
   ||| ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : :
Db 984 ATGTGCATGGGAGGAGACAGTCTGTAACTCGGGGACAGACCTCTTTGAAGACATA 1043

QY 232 ---AlaThrPheSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSer 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1044 GAAAGCTACATAGTTGTGATGATGCCCTCACCAGTCCGAGAGACAGATCATCTTTCAC 1103

QY 251 ThrValProLys-----ThrGlnAsp----- 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1104 AGCTGCAGCGGTTCCTCCCTCATGACAGTTCGACATCAGAGACATATCCGGGCGCTG 1163

QY 258 -----GlnArgProGlnLeuAsp 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1164 CCCAGCCACAGCCACAGGACCTTCAGAGCATATGTCGGGACGCTGTGACATCAAT 1223

QY 264 ProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIle 283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1224 AAGGACAAAGCTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1283

QY 284 GlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAsnLysIleThr 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1284 CAGCTATGCGAGTTCCTCTGGAATTAATCTGATTAATCTGATCTTTATCAGC 1343
QY 304 TrpGlnGlyThrAsnGlnGlyIlePheLysMetThrAspProAspGlnValAlaArgTrp 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1344 TGGACAGGAGATGGCTGGGAATTCACCTTCTGACCCAGATGAGTGGCCAGAGATGG 1403

QY 324 GlyIleArgLysSerLysProAsnMetAsnThrLysAspLysLeuSerAlaLeuArgTyr 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1404 GGAAGAGGAAAAACAACCTAGATCAATTAATGAGAACTGAGCGCTGAGCTACGCTAC 1463

QY 344 TyrTrpAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPhe 362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1464 TATTACGACAAAACATCATCCACAGACAGCGGAGAAAGCTAAGCTGATCCGCTTGTG 1523

QY 363 ---AspPheHisGlyIleAlaGlnIleAlaLeuGlnProHisProGlu 377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1524 TGTGACCTG-----CAGAGCTCTGGGGTACACCCCTGAG 1559

RESULT 8
US-09-344-579-1
; Sequence 1, Application US/09344579
; Patent No. 6054316
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Lex M. Cowert
; TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
; FILE REFERENCE: RTS-0063
; CURRENT APPLICATION NUMBER: US/09/344, 579
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (291)..(1700)
US-09-344-579-1

Alignment Scores:
Pred. No.: 8.63e-37 Length: 2268
Score: 445.50 Matches: 127
Percent Similarity: 40.96% Conservative: 43
Best Local Similarity: 30.60% Mismatches: 105
Query Match: 18.23% Indels: 141
DB: 3 Gaps: 14

US-09-902-772-2 (1-451) x US-09-344-579-1 (1-2268)
QY 122 GlnArgArgValIleValIProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 CACGGCGCTGGGCGCATCCAAAGAACCCCTGGCTGTGAGTGAAGCAACAGGTATGCCAG 629

QY 142 TrpLeuGlnTrpAlaValIleGlnGlyLysLeuProAspValAlaPheLeuPheGln 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 TGGCTTCTCTGGGCGCACCATGATGATGATGATGATGATGATGATGATGATGATGAT 666

QY 162 AsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 687 GGCATGAATGGCAGATGCTGTGTAACCTTGGCAAGAGAGCGTTTCTGGAGCTGGCAGCT 746

QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisLysIleValArgGlnArgGlyAlaThr 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 747 GACTTGTGGGTGATCTCTCTGGGAAACATCTGAGCAAAATGATCAAGAAAAACAA--- 803

QY 202 PheIlePheProAsnThrSerValIleProGlnAlaIleThrGlnArgIleThrArgPro 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 804 -----GAAAGACAGAGATCAATATGAAGAAATTCACACCTCAGCTCGTTCCT 854

QY 222 -----AspLeuProTyr-GlnGlnAl 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 CATTGATTAAACAGCAATACATTAGGTTTGGCACAGAGAGCGCCCTATGGAATGACAG 914

```



```

Db          427 TCACGTACAAGTTCATTTCACAAACAGTGTCTGTGTCATTCACATTCATTGATGTGG 468
Oy          366 -----GlytIleAlaGlnAlaLeuGlnProHisProGlnSerMet 381
Db          487 GCTTGCGTGGGGGTGCAGTCCAGAGTGCCTCCGCGAGTCCGCGGTGGTGGCCACT 546
Oy          381 ytlstlyrProSerArleProTylMetSerSetTylHisAlaHisProGlnMet 401
Db          547 TCGGCTTCCTCCCTTCACGCC-----TCG 573
Oy          401 snheValAlaPro-----HisProAlaLeuProValrhesers 416
Db          574 AGGTGCGTGTCCCGCCAGGAGACCCCGCTGCACGACGCGTCTTCATTCATTCATT 633
Oy          416 erphepheaIaIa 420
Db          634 CCTCTCTCTCGCT 647

RESULT 10
US-09-021-715-1
; Sequence 1, Application US/09021715
; Patent No. 6194547
; GENERAL INFORMATION:
; APPLICANT: MayoThalassitis, George J.
;           Blair, Donald G.
;           Fisher, Robert J.
;           Beal Jr., Gregory J.
;           Athanasiou, Meropi A.
;           Sgouras, Dionysios N.
; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3634
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,715
; FILING DATE: 10-Feb-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Mackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 015280-229000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2667 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 123..1769
; OTHER INFORMATION: /note= "human ERF (ETS2 Repressor
; Factor) cDNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-021-715-1

Alignment Scores:
Score: No.: 7 49e-26 Length: 2667
Score: 342.50 Matches: 84
Percent Similarity: 53.408 Conservative: 26

```

```

Best Local Similarity: 40.78%  Mismatches: 59
Query Match: 14.01%  Indels: 38
DB: 4  Gaps: 5

US-09-902-772-2 (1-451) x US-09-021-715-1 (1-2667)

QY 239 ProthrgInserLysAlahrcInProserSerSerThValProLysrThrgInuasp--G 258
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 74 CCGGAGCGCGCGCGCGCGGAGTGGGGCGCTGGCCCCCGGGCCCCCGACGTGAAGACCCC 133
    258 InarProGInleuAaspProTyrgInleuGlyProThrSerSerArgLeuAlaAsp 278
    134 GCGGACACAGGGGTTTGCTTCCCGGAT--GGGCTACA--ACCGAG-TCGTCCC 166
    278 roGlySerGlyCInleGInleuTrpGInPheLeuLeuGInleuLeuSerAspSerSera 298
    187 CTGGCTCAAGGACAGATCCAGCTGTGGCACTTATATCTGGAGGCTGCGGAGAGAGT 246
    298 snserAsnGylleThrTrpGInuGlyThrAsnGlyInPheLysMetThrAspProaspG 318
    DB 247 ACCAGGCGCTCAATTCCTGGCAGGGGAGCTACGGGGAATTCGTATCAAGACCTGATG 306
    318 luValAlaArgrArgrTrpGlyInuArgLysSerLysProAsnMetAsnTyAspLysLeus 338
    307 AGTGCGCCCGCGCTGTGGGGCGCTTCCGAATGCAGAGCCAGAGATTAAGACAGCTGA 366
    QY 338 eArGAlaLeuArgrTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 356
    DB 367 GCGGGCGCTGTGTATTAATTAACCAACGCGATTCGTGACAAAGCAAGGGAACGGT 426
    QY 358 yTAlArTyLysLysheAspPheHis----- 365
    DB 427 TCACCTCAAGATTCATTAATTCACAACTAGTGTGTCAATTACCATTCATTTGATGTGG 486
    QY 366 -----GlyleAlaGInAlaLeuGInProHisProGInuSerSerMet 381
    DB 487 GGTGGCTGGGGGTGCAGTGCACAGAGTCCCGCCGACAGTCCGTGGGGGTGAAGCAGT 546
    QY 381 yTLysTyTyProSerAspLeuProTyTyMetSerSerTyHisAlaHisProGInLysMet 401
    DB 547 TCCGCTCCCTCCCTCCCAAGCGCC-----TCCG 573
    QY 401 snPheValAlaPro-----HisProAlaLeuProValThrSerSers 416
    DB 574 AGGTCTGTGCCCCACGAGAGACCCCGCTCAACACAGCCTGCTTCATCTTCATCTT 633
    QY 416 erPhePheAlaAla 420
    DB 634 CCTCTCTCGGCT 647

RESULT 11
US-08-875-944B-1
; Sequence 1, Application US/08875944B
; Patent No. 6036542
; GENERAL INFORMATION:
; APPLICANT: FUJINAGA, Kei
; APPLICANT: YOSHIDA, Koichi
; APPLICANT: HIGASHINO, Fumihito
; TITLE OF INVENTION: CANCER CONTROL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```



```

; LENGTH: 2410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-780-835B-1

```

Pred. No.:	4.35e-22	Length:	2410
Score:	306.00	Matches:	109
Percent Similarity:	36.67%	Conservative:	34
Best Local Similarity:	27.95%	Mismatches:	107
Query Match:	12.52%	Indels:	141
DB:	2	Gaps:	14

US-09-902-772-2 (1-451) x US-08-780-835B-1 (1-2410)

QY	45	SeRysmEtSerProAtgVAlPrOGInLnsPrLSeuSerG1PrProAlArVAl	64
Db	634	TCAGAAAGAAACCCAGAGTCCC-----GCACACACCCCGCCGTCT	678
QY	65	Thr1LeuYmEtGlucYAsnProAnGlnValAsnGlySerArGAsnSerProAsp	83
Db	679	GCACAGGAAGACCACTCCCTACCAACACATGGAGACAGTCCCTTACTACCAACAA	738
QY	84	-----AspCySerValAlAsy	89
Db	739	TCCGCATCAGTCCCGGCTCCGGTCCCTGGACAGTCCGCTCGAGCCCTTTTCCA	798
QY	90	GlYGlYLySmEtVAlSerSerSerAspAsnValGlYmEtAsnTYGlYSerTYmEtGln	109
Db	799	GGGAGAAACAGACAGACAGCTCC-----TGAGAG	828
QY	110	GlYdYsHs1LeProProProAsmEtHrThrAsnGlnAdgArVAl1LeVAlProAla	129
Db	829	CTTCACAGCTCTCCAGTCCCACTCCCGCCAGCGGTACCTTGAGACAGACTCCGT	888
QY	130	AspProThrLeuTrpSerThrAspHisValArGlnTrLeuGlnTrPAlaValYsGln	149
Db	889	TTCAGCAGCCCGTGGACA-----	906
QY	150	TyrGlYLeuProAspValAsp1LeuLeuPheGlnAsn1LeAspGlyLysGluLeuGys	169
Db	907	-----TGT	909
QY	170	LySmEtHrLyAspAspPheGlnArGLeuHrProSerTYrAsnAlAsp1LeuLeu	189
Db	910	-----GCCACTCCCTCA-----	921
QY	190	SeRHisLeuHs1TYrLeuArG1uArG1uAlArHrPhe1LePro-----AsnThr	207
Db	922	-----CATCTCTCAGGGAGGGGGCGGGAACCTTCGACGCCCATCAACAC	972
QY	208	SeR-----ValTYrProGlnAlArHrGlnArG-----	216
Db	973	ACSTGTGGAGCCCTGCCACCCCTACCCCGCCAGCAATTCAAGAGAGTACCATGACC	1033
QY	217	---1LeHrHrArGrProAspLeuProTYrGluGlnAlArGrSerAla-----	232
Db	1033	CCCTGTACGAACAGAGCTGGCCAGCCGCTT---CAAGCCAGGGTGGGTCACTAGTGGCACA	1086
QY	233	-----TrpHrSerHisSerHis1ProHrGlnSer1YsAlArHrGln	246
Db	1090	GGTACCCAGAGGGGGGTGTGATCAACAGAGCGCAGACTTGCCTACAGACTGAG	1144
QY	247	ProSerSerSerHrValProLYsThrGlnAsp1nArProGln1LeuAspProTYrGln	266
Db	1150	ATGTCTCCCTGATGTGCATCAATGTACTCCACCCAGAGGGCTCTCTGGACCTTCACG	1209
QY	267	1Le-----	267
Db	1210	GTGATGAGAGTATGGGTATGGCTATGAAAATCCCTTGACACATTCGCCAGATGATGCT	1266
QY	268	-----LeuGlyProHrSerSerArG-----LeuAlAsn	277

[illegible]

```

RESULT 14
US-09-303-268-1
: Sequence 1, Application US/09303268
: Patent No. 6172212
:
: GENERAL INFORMATION:
: APPLICANT: Hung, Mien-Chie
:           Xing, Xiangming
: TITLE OF INVENTION: PEAS is a Tumor Suppressor
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ARNOLD, WHITE AND DURKEE
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77210-4433
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/303,268
: FILING DATE: 30-Apr-1999
: CLASSIFICATION: <Unknown>
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/780,835
: FILING DATE: 10-JAN-1997
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wilson, Mark B.
: REGISTRATION NUMBER: 37,259
: REFERENCE/DOCKET NUMBER: UTSC500
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (512) 474-7577
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 2410 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-303-268-1
:
: Alignment Scores:
: Pred. No.: 4.35e-22      Length: 2410
: Score: 306.00           Matches: 109
: Percent Similarity: 36.67% Conservative: 34
: Best Local Similarity: 27.95% Mismatches: 107
: Query Match: 12.55%      Indels: 141

```

DB:	4	gaps:	14
US-09-902-772-2 (1-451) x US-09-303-268-1 (1-2410)			
Oy	45 SerIysMetSerProArgValProGlnGlnAspIleuSerGlnProProAlaArgVal	64	
Db	634 TCAGAAAGAAAGAACCCACAGATCCC-----GCACGAGACCCGCCCTTCCT	678	
Oy	65 ThrIleIysMetGluCysAsnProAsnGlnValAsnGlySerAraGlnSerProAsp---	83	
Db	679 GCACGAGAAACCCACACTCCCTTACACCATGTGAGAGACACTGGCTTTACTCCAGACAAA	738	
Oy	84 -----	89	
Db	739 TCGCCATCAAGTCCCGCGCTCCCGGTGCCGTGGACAGTCGCCCTCCGAGCCCTTTTCA	798	
Oy	90 GlnGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySerTyrMetGlu	109	
Db	799 GGGCAGAAAGACAGACAGACCTCC-----TGAGAG	828	
Oy	110 GlnIysHisIleProProProAsnMetThrThrAsnGlnIleArgArgValIleValProAla	129	
Db	829 CCTCAGAGCTTCCACAGTCCACACCTCGGACGGGTACCTGGTGGAGACAGCTCCGCT	888	
Oy	130 AsproThrIleuTyrSerThrAspHisValArgGlnIleuGlnIleuTyrPalaValLysGlu	149	
Db	889 TCCAGCAGCCCGTGAGCA-----	906	
Oy	150 TyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLysGluLeuCys	169	
Db	907 -----TGT	909	
Oy	170 LysMetThrIlyAspAspPheGlnArgLeuThrProSerTyrAsnAlaAspIleLeuLeu	189	
Db	910 -----GCCACTCCTTCA	921	
Oy	190 SerHisIleuHisTyrLeuArgGlnArgGlyValAlaThrPheIlePhePro-----AsnThr	207	
Db	922 -----CAVTCCTCAGGAGGAGGAGGCGCGGAGCACTCTCCACGCCCTTCATCAACAC	972	
Oy	208 Ser-----ValTyrProGlnAlaThrGlnArg	216	
Db	973 AACGCTGGAGACCTCGCCACCCATCCACCCAGCAGAACCTTCAGCAGAGTACATGAC	1032	
Oy	217 ---IleThrThrArgProAspLeuProTyrGlnIleGlnAlaArgArgSerAla-----	232	
Db	1033 CCCGTAGACAGACAGCTGCCACGCCGCT---CAAGCAGAGGTGGGTCACTAGTGGCACA	1089	
Oy	233 -----TyrThrSerHisSerHisProThrGlnSerLysAlaThrGln	246	
Db	1090 GGTACCCAGGCGGGGTGGTGTATGAACAGCAGCAGCAGACACTTCGCGCTACACATCAG	1149	
Oy	247 ProSerSerSerThrValProLysThrGlnAspGlnArgProGlnIleuAspProTyrGln	266	
Db	1150 ATGTCCTCGATGTGCATCAATGTACCTCCACCCAGAGGCTTCTCTGGACCTTCACG	1209	
Oy	267 Ile-----	267	
Db	1210 GTGATGAGAGTATGGTATGATAAATCCCTTCGACCATTCACAGATGATGCT	1269	
Oy	268 -----LeuGlyProThrSerSerArg-----LeuAlaAsn	277	
Db	1270 GCATTGTCCCTTAAATAATTTGAGAGACAGACAGAGAGGATGTGAGACTTTCGGG	1329	
Oy	278 ProGly-----SerGlyGlnIleGlnIleuTyrPheIleuLeuGlnIleu	292	
Db	1330 AG-GGGCGACCCCTTACACGCGCGGGGTGCTTCAACACTGTGCGAGTTCGTGGGCCCTG	1388	
Oy	293 LeuSerAspSerSerAsnSerAsnCysIleThrTyrPheGlnIleThrAsnGlyLysPheLys	312	
Db	1389 CTGATGATGCCACAAATATGCTCATTTCACTTCTCTGGACAGGCGCGGAGATGAGATTAA	1448	
Oy	313 MethrAspProAspGlnValAlaIleArgArgTyrPheGlyGlnArgLysSerLysProAsnMet	332	

```

Db      1449 CTAATTGAACCTGAAAGAGTTCCAGCGCTCTGGGGATTCACAGAAGACC GCCAGCATG 150
        :::: :|||||:||||||| ||||| :|||:||||:|||| |
Qy      333 AsnTyrAspLysLeuSerArgAlaLeuArgTrpTyrTyrAsnIleMetThrLys 352
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1509 AATTATGACAACCTGACGCCGTCCGCTGCATACTATTATGACAAAGCATTATGCAGAAG 156
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      353 ValHisGlyLysArgTyrAlaTyrLysPhe 362
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1569 GTGCTGGCGAACGCTACGTGTACAGATT 1598
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-09-116-049-1
; Sequence 1, Application US/09116049A
; Patent No. 6248351
; GENERAL INFORMATION:
; APPLICANT: Hung, Men-Chie
; TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
; FILE REFERENCE: UTSC:582
; CURRENT APPLICATION NUMBER: US/09/116,049A
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2410
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-116-049-1

Alignment Scores:
Pred. No.:          4,35e-22           Length:         2410
Score:              306.00             Matches:          109
Percent Similarity: 36.67%             Conservative:     34
Best Local Similarity: 27.95%           Mismatches:       107
Query Match:        12.52%              Indels:           141
DB:                 4                  Gaps:            14

US-09-902-772-2 (1-451) x US-09-116-049-1 (1-2410)
Qy      45 SerLysMetSerProArgValProGlnGlnAsnPTripLeuSerGlnProProAlaArgVal 64
        |||::: :|||||:||||||| ||||| :|||:|||| |
Db      634 TCATAAGAGAACCCCAAGAGCTCCC-----GCACAGACCCCCCGCTCTCCT 678
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      65 ThrIleLysMetGluCysAnpProAsnGlnValAsnGlySerArgAsnSerProasp--- 83
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      679 GCAGCAGGAAGCCACACACTCCCTTACCACATGAGAGAGAGTGCCCTTACTCCAGACAAA 738
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      84 -----AspCysSerValAlaLys 89
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      739 TCGCCATCAAGTCCCCCGCTCCGGTCCCGCCCTGCAGACAGTGC CCCCTGCAGCCCTTTCCA 798
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      90 GluGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrcIyserTyrmetGlu 109
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      799 GGCGCAAGACGACGACGAGACCTCC-----TGAGAG 828
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      110 GluLysHisIleProProProAsnMetThrThrAsnGluArgArgValIleValProAla 129
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      829 CCTCCAGCGCTTCCAGTCCACAGTCCACCTGGGCCACGCGGTGATGACACAGCTCCGCTC 888
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      130 AspProThrIleTrpSerThrAspHisValAlaArgGlnTrpLeuGlnUrrPalalaValLysGlu 149
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      889 TCCACAGCACCCCTGGACA----- 906
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      150 TyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspLysGluLeuLys 169
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      907 -----TGT 909
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      170 LysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAspIleLeuLeu 189
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      910 -----GCCACTCCGTCA----- 921
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      190 SerHisLeuHisTyrLeuArgGluArgGlyAlaThrPheIlePhePro-----AsnThr 207
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 922 -----CATCTCTCAGGAGGCGGCGGAGACCTTCCAGCCCCCTATCAACACC 972
QY 208 Ser-----ValTYrProGluAlaThrGlnArg----- 216
Db 973 AACGTGGAGCCCTGCCACCCCTACCCCGCAGAACTTCAAGCAGAGTACCATGACC 1032
QY 217 ---lIeThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAla----- 232
Db 1033 CCCTGTACGAAACAGCGTGGCCAGCCGCTT---CAAGCCAGGGTGGGGTCTAGTGGCACA 1089
QY 233 -----TTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln 246
Db 1090 GGATCCAGAGGGCGGGGGGTGGATCAACAAGAGAGCCACAGACTTCGCTACGACTCAG 1149
QY 247 ProSerSerSerThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGln 266
Db 1150 ATGTCCCTGGATGTGCATCATATGATACCTCCACCGAGAGGGCTTCTGTGACCTCTCCAG 1209
QY 267 lIe----- 267
Db 1210 GTGATGGAGTGTGTTATGCTATGAAAAATCCCTTCGACATTCACAGATGATGCT 1269
QY 268 -----LeuGlyProThrSerSerArg-----LeuAlaAsn 277
Db 1270 GCATTGTCCCTAAAAAATTGGAAGAGACATCAAGCAGAGAGGATTGGAGCTTCCGGG 1329
QY 278 ProGly-----SerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeu 292
Db 1330 AG-GGGCCACCCCTACAGCGCGGGGTGCTTACACTGTGGCACTTGTGGCCCTG 1388
QY 293 LeuSerAspSerSerAsnSerAsnCyslIeThrTrpGlnGlnIleThrAsnGlyGluPheLys 312
Db 1389 CTGGATGACCCCAACAATGCTCATTTCTGCTTGGACAGCGCGGGGATGAGATTAA 1448
QY 313 MetThrAspProAspGluValAlaArgArgTrpGlyLysGlySerLysProAsnMet 332
Db 1449 CTAATTTAACCCTGAAGAGTTGCCGCTCTGGGGTATCCAGAAGAACCGCCAGCATG 1508
QY 333 AsnTYrAspLysLeuSerArgAlaLeuArgTYrTYrAspLysAsnIleMetThrLys 352
Db 1509 AATTATGACAAAGCTGACCGCTGCTGCTGATCTATATGAGAAAGCATCATCAGAG 1568
QY 353 ValHisGlyLysArgTYrAlaTYrLysPhe 362
Db 1569 GTGGCTGGCAGACGCTAGTGTACAACTT 1598
```

Search completed: July 28, 2003, 07:41:33
Job time : 76.6835 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 06:34:34 ; Search time 284.484 Seconds

(without alignments) : 3270.525 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444

Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPNTRLPAAHMPHSLCTYY 451

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1439767 segs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL-frame_p2n_model -DEV-xlh
-O-/cgn2_1/USPTO.spool/US09902772/funat_23072003_093658_14962/app_query.fasta_1.1294
-DB-published.Applications_NA -OFMT-fastap -SUFFIX-rnmp -MINMATCH=0.1
-LOOCL-0 -LOOEXT-0 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MTLEN=0
-MAXLEN=2000000000 -USER=US09902772@cgn_1_1.82_erunat_23072003_093658_14962
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSELOCK=100
-LONLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PTI_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
11:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2444	100.0	1447	11	US-09-902-772-1 Sequence 1, Appl1

2	2414.5	98.8	1528	11	US-09-902-772-3	Sequence 3, Appl1
3	2182	89.3	3166	15	US-10-205-823-98	Sequence 98, Appl1
4	948	38.8	567	10	US-09-864-761-18410	Sequence 18410, A
5	863	35.3	473	10	US-09-864-761-20472	Sequence 20472, A
6	712	29.1	549	10	US-09-923-779-54	Sequence 54, Appl1
7	650	26.6	420	10	US-09-864-761-3705	Sequence 3705, Ap
8	650	26.6	454	10	US-09-864-761-1653	Sequence 1653, Ap
9	638	26.1	467	12	US-09-918-995-23366	Sequence 23366, A
10	511	20.9	1752	10	US-09-850-799-1	Sequence 1, Appl1
11	445.5	18.2	2269	11	US-09-954-531-955	Sequence 955, App
12	444.5	18.2	2188	11	US-09-920-300A-1716	Sequence 1716, Ap
13	444.5	18.2	2188	15	US-10-033-528-1716	Sequence 1716, Ap
14	444.5	18.2	2268	11	US-09-920-300A-1693	Sequence 1693, Ap
15	444.5	18.2	2268	15	US-10-033-528-1693	Sequence 1693, Ap
16	444.5	18.2	3692	15	US-10-106-698-824	Sequence 824, App
17	431.5	17.7	1884	11	US-09-925-300-420	Sequence 420, App
18	407	16.7	2532	15	US-10-037-270-869	Sequence 869, App
19	374	15.3	553	10	US-09-864-761-7427	Sequence 7427, Ap
20	371	15.2	225	10	US-09-864-761-21761	Sequence 21761, A
21	371	15.2	348	10	US-09-864-761-29364	Sequence 29364, A
22	371	15.2	477	10	US-09-864-761-5035	Sequence 5035, Ap
23	371	15.2	519	10	US-09-864-761-12798	Sequence 12798, A
24	339	13.9	3178	15	US-10-108-605-124	Sequence 124, App
25	339	13.9	3178	15	US-10-108-605-128	Sequence 128, App
26	331	13.5	473	12	US-09-918-995-27904	Sequence 27904, A
27	322	13.2	2064	12	US-09-884-363-3	Sequence 3, Appl1
28	322	13.2	2064	15	US-10-167-485-1	Sequence 1, Appl1
29	322	13.2	2333	11	US-09-920-300A-1788	Sequence 1788, Ap
30	322	13.2	2333	11	US-09-880-107-3316	Sequence 3316, Ap
31	322	13.2	2333	15	US-10-033-528-1788	Sequence 1788, Ap
32	306	12.5	2410	12	US-09-884-363-1	Sequence 1, Appl1
33	303	12.4	416	12	US-09-918-995-8472	Sequence 8472, Ap
34	301	12.3	1894	10	US-09-126-945B-1	Sequence 1, Appl1
35	301	12.3	1905	10	US-09-866-356-2	Sequence 2, Appl1
36	301	12.3	1942	15	US-10-157-931-83	Sequence 83, Appl1
37	300.5	12.3	2952	10	US-09-955-297-257	Sequence 257, App
38	300.5	12.3	4106	15	US-10-084-817-187	Sequence 187, App
39	296	12.1	300	15	US-10-177-063-15	Sequence 15, Appl
40	287	11.7	2256	15	US-10-098-841-317	Sequence 317, App
41	282	11.5	1087	15	US-10-102-806-215	Sequence 215, App
42	281	11.5	300	15	US-10-177-063-16	Sequence 16, Appl
43	280	11.5	1894	10	US-09-841-963A-1	Sequence 1, Appl1
44	279	11.4	165	10	US-09-864-761-23294	Sequence 23294, A
c	260.5	10.7	1429	11	US-09-764-864-320	Sequence 320, App

ALIGNMENTS

RESULT 1
US-09-902-772-1
Sequence 1, Application US/09902772
Patent No. US20020164739A1
GENERAL INFORMATION:
APPLICANT: Chuai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
TITLE OF INVENTION: the Proteins
FILE REFERENCE: chuai selyaku kabushiki kaisha 5001
CURRENT APPLICATION NUMBER: US/09/902,772
PRIOR APPLICATION NUMBER: 2001-07-12
PRIOR FILING DATE: 1997-06-18
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1447
TYPE: DNA
ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
US-09-902-772-1
Alignment Scores:
Pred. No.: 1.09e-256 Length: 1447
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 100.008 Indels: 0
 DB: 11 Gaps: 0

US-09-902-772-2 (1-451) x US-09-902-772-1 (1-1447)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 63 ATGGCAAGCACTATTAGAGACATATATCATGTGTGATGAGACACGCTCTTTGAG 122
 QY 21 CysAlaTyrGlnSerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 123 TGTGCTACGAGATGCCCGACCTTGGCAAAAGCAAAATGACAGCTCTCTCCAGTGA 182
 QY 41 TyrGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 DB 183 TATGGCAAAACATCAAGATGAGCCCGCGGTGCCAGCAGAGCTGTTATCACACCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 243 CCGGCCAGAGTTACATTAGATGAGAGTAAACCAACAGGTTAATGGTCAAGGAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 303 TCACCTGATGACTGACGCGTGGCAAAAGAGGAAATGGTTAGCACTTCAGACAAATGTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThrThr 120
 DB 363 GGGATGAACATATGGAACCTACATGGAAGAGAGATATCCGCTCCAAATATGCAAC 422
 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleutrpSerThrAspHisValArg 140
 DB 423 AATGAAGAAAGATTATGTGCCAGCAGATCTACGTTATGAGACACAGCATGTACGG 482
 QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 DB 483 CAGTGGCTGAGTGGGAGAGAGAGATGATGCTTCCAGACGCGGACATCTGTGTTTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 543 CAGAACATGATGGGAAAGAGTGTGTAAATGACCAAGATGACTCCAGAGAGCTACG 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAla 200
 DB 603 CCGAGCTATACGACGATATCCCTGTCACACCTACACCTCAAGAGAGAGAGGCC 662
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrArg 220
 DB 663 ACTTTATTTTCCAAATATACATCATGTTTACCCAGAGACGCAAAAGATTAACACAGG 722
 QY 221 ProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrpHisSerHisSerProThr 240
 DB 723 CCAGATTACCTTATGAGCAAGGAGAGATCAGCGAGATCCAGCATCCCACT 782
 QY 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgPro 260
 DB 783 CAGTCAAAAGCTACCAACCATCATCTCAACAGTGGCCAAACAGAGACAGCGCTCT 842
 QY 261 GlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
 DB 843 CAGTTGATCTTATCAGATTCTTGGACGACAGCAGCGCTCTGCAAAATCCAGGAGT 902
 QY 281 GlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuSerAspSerSerAsnSerAsn 300
 DB 903 GGGCAGATACAGCTATGCGAGTCTCACTGAGACTTCTGCGAGAGCTCCAACTCCAC 962
 QY 301 CysIleIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThrAspProAspGluValAla 320
 DB 963 TGGATCACCTGGAGGAGCAAAATGGGAGTTCAAGATGACAGACCTGATGAAGTGGCT 1022
 QY 321 ArgArgTrpGlyLysGlySerLysProAsnMetAsnTyrAspLysLeuSerArgAla 340
 DB 1023 CGGCGTTGGGAGAGAGAAACCAACCTAACATGACTATGACAAACTCAGCCGTGCA 1082

QY 341 LeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTyr 360
 DB 1083 CTTCGCTACTACTATGACAAAAATATATATGACTAACTGATGTAACGCTATGCTTAC 1142
 QY 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerMet 380
 DB 1143 AAATTTATTTCCAGGAAATCGCTCAGGCCCTCCAGCTCACTCCAGATATCATCAT 1202
 QY 381 TyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLysMet 400
 DB 1203 TACAATACCATATAGACCTCCCTACATGATGTTCTTACCATGACACCCCGAGAAATG 1262
 QY 401 AsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPheAlaAla 420
 DB 1263 AACTGTAGCTCCCATCCCGCGCTTGGCGGTATACATCATCAGCTTTTGTCTGCC 1322
 QY 421 ProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrProAsnThrArgLeuProAla 440
 DB 1323 CCTATATCATATGGAATTCACCAACTGAGGAGCATCTACCCCAATACAGCTGCCAGCT 1382
 QY 441 AlaHisMetProSerHisLeuGlyThrTyrTyr 451
 DB 1383 GCTCATATGCTTCCATCTTGGCAGCTACTAC 1415

RESULT 2
 US-09-902-772-3
 ; Sequence 3, Application US/09902772
 ; Patent No. US20020164739A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
 ; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
 ; FILE REFERENCE: Chugai seiyaku kabushiki kaisha 5001
 ; CURRENT APPLICATION NUMBER: US/09/902,772
 ; PRIOR FILING DATE: 2001-07-12
 ; PRIOR APPLICATION NUMBER: US/08/878,177
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1528
 ; TYPE: DNA
 ; ORGANISM: c-erg gene, chicken DNA
 ; US-09-902-772-3

Alignment Scores:
 Pred. No.: 1,94e-253 Length: 1528
 Score: 2414.50 Matches: 450
 Percent Similarity: 94.14% Conservative: 0
 Best Local Similarity: 94.14% Mismatches: 1
 Query Match: 98.79% Indels: 27
 DB: 11 Gaps: 1

US-09-902-772-2 (1-451) x US-09-902-772-3 (1-1528)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 63 ATGGCAAGCACTATTAGAGACATATATCATGTGTGATGAGACACGCTCTTTGAG 122
 QY 21 CysAlaTyrGlnSerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 123 TGTGCTACGAGATGCCCGACCTTGGCAAAAGCAAAATGACAGCTCTCTCCAGTGA 182
 QY 41 TyrGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 DB 183 TATGGCAAAACATCAAGATGAGCCCGCGGTGCCAGCAGAGCTGTTATCACACCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 243 CCGGCCAGAGTTACATTAGATGAGAGTAAACCAACAGGTTAATGGTCAAGGAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100

```

Db      303 TCACCTGATGACTGCAGCGGTGGCAAAAGAGGAAATGCTTAGCAGATTCAACAAATGTT 362
Qy      101 GYMETASNTYGLYSerTYrMetGLUGLYSHISLIErProProAsmMetThrThr 120
Db      363 GGGATGAACCTATGGAAGCTACATGGAAGAAGCAATATTCGCTCCAAATAATGCAACC 422
Qy      121 ASNGUATARGVALILEVALProAlaAspProThrLeuTrpSerThrAspHisValArg 140
Db      423 AATGACGAAGAAGTATTGTGTGCACACATCTCTACGTTATGAGACACGACCATATGAGG 482
Qy      141 GlnTrpLeuGLUTrPAlaValIlySGlUTrYrGlyLeuProAspValAspIleLeuLeuPhe 160
Db      483 CAGTGGCTGGAGTGGGAGTGAAGAGATGAGTGTTCACAGAGCTGACATCTGTTGTTC 542
Qy      161 GlnAsnIleAspGlyLySGlUleucYslySMetThrLySAspAspPheGlnArgLeuThr 180
Db      543 CAGAACATTGATGGGAAAGAGTGTGTAATAATGACCAAAAGATGACTTCCAGAGACTCAGC 602
Qy      181 ProSerTYrAsnAlaAspIleLeuLeuSerHisLeuHisTYrLeuArgGluArg----- 198
Db      603 CCGAGCTTAAAGCAGATATCTCTCTGTACACACTACACTACCTTAGAGAGAGACTCTCTT 662
Qy      198 ----- 198
Db      663 CCACATTGACTTCAGATGATGTTGATAGGCTTACAAAACCTCCACGGTTAATGCAT 722
Qy      199 -----GlyAlaThrPheIlePheProAsnThrSerValTYrProGluAla 213
Db      723 GCTGAAACACAGAGAGGACCACTTTATTTTTCAAATACATCAGTTTACCAGAAACA 782
Qy      214 ThrGlnArgIleThrThrArgProAspLeuProTYrGlnGlnAlaArgSerAlaTrp 233
Db      783 ACGCAAAATAATACAAAGAGCCAGATTACCTTATGAGCAAGCAGAGAGATCAGCGTGG 842
Qy      234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValPro 253
Db      843 ACGAGTCAACAGCATCCACTCAGTCAAGCAAAAGCTAACCCACATCATCTTCAACAGTCCC 902
Qy      254 LysThrGluAspGlnArgProGlnLeuAspProTYrGlnIleLeuGlyProThrSerSer 273
Db      903 AAAACAGAGAGACAGCGCTCTCAGCTTACATCTTATCAGATTCTTGACCCAGCAGCAGC 962
Qy      274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuLeu 293
Db      963 CGTCTTGAATTCAGAGGAGAGTGGGAGATACAGCTATGCGAGTTCCTTACGAGCTTCTG 1022
Qy      294 SerAspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMet 313
Db      1023 TCGGACAGCTCCAACTCCAACTGCATCCTGGGAGGSCAAATGGGGAGTTTCAAGATG 1082
Qy      314 ThrAspProAspGluValAlaIleArgArgTrpGlyGluArgLysSerLysProAsnMetAsn 333
Db      1083 ACAAGCCCTGATGTAAGTGGCTGGCGTTGGGAGAGAGGAAACCAAACTAACATCATGAC 1142
Qy      334 TyrAspLysIleuSerArgAlaLeuArgTYrTYrTyrAspLysAsnIleMetThrLysVal 353
Db      1143 TATACAAACTCAGCCGCTGCTGCTACTACTATGACAAATAATTATATACATPAAGTT 1202
Qy      354 HisGlyLysArgTYrAlaTYrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
Db      1203 CATGCTAAACGCTATGCTTACAAATTTGATTCCACGGAATCGCTCAGGCCCTCAGACCT 1262
Qy      374 HisProProGluSerSerMetTYrLysTYrProSerAspLeuProTYrMetSerSerTYr 393
Db      1263 CACCTCTCAGATATCATCATGTATCAAAATACCCATCAGACCTCCCTCATATGATGTTCTTAC 1322
Qy      394 HisAlaHisProGlnLysMetAsnPheValAlaIleProHisProProAlaLeuProValPhe 413
Db      1323 CATGCACACCCCGAAGATGAACTTTGTAGCTCCCATCCCTGCTTGGCCCGTAAACC 1382
Qy      414 SerSerSerPhePheAlaIleProAsnProTYrTrpAsnSerProThrGlyGlyIleTYr 433

```

```

Db      1383 TCATCCAGCTTTTGTGTCGCCCTTAATCCATACCTGGAATTCACCAAGTGGAGCATCTAC 1442
Qy      434 ProAsnThrArgLeuProAlaIleHisMetProSerHisLeuGlyThrTYrTYr 451
Db      1443 CCCAATACCAGGCTCCAGCTGCTCATATGCTTCCCATCTTGGCAGCCTACTAC 1496

RESULT 3
US-10-205-823-98
; Sequence 98, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gordatcheva, Bella
; APPLICANT: Hoersck, Sebastian
; APPLICANT: Kamalakar, Shubhangt
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatf, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-98.

Alignment Scores:
Pred. No.: 1,27e-227 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 15 Gaps: 4

US-09-902-772-2 (1-451) x US-10-205-823-98 (1-3166)
Qy      2 AlaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db      281 GCACCTCATATCAAGGAAGCCTTATCAGTTGTGAGTGAGGACCAAGTCTGTGAGTGT 340
Qy      22 AlaTYrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerSerGluTYr 41
Db      341 GCCTACGGAACGCCACACCTGGCTPAAGACAGAGATGACCGGCTCTCTCCAGGAGACTAT 400
Qy      42 GlynInThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db      401 GGACAGACTTCCAGAGTGAAGGCCACGGCTCCTCAGCAGATTTGGCTGTCTCAACCCCA 460
Qy      62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db      461 GCCAGGGTCAACCAATGAATGGAATGTAAACCTTGAACCGTGAAGTGAATGCTCAAGAACTCT 520
Qy      82 ProAspAspCysSerSerValAlaLysGlyGlyLysMetValSerSerSerSerAspAsnValGly 101

```

Db	521	CCTATGATGATCGACGTGGCCCAAGGCGGGAAGATGTTGGGACGACCCAGACACCGTTGGG	580
OY	102	MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrAsn	121
Db	581	ATGAACACTCGGACAGCTACATGAGAGAGAACATGACCCACCCCAACATACCCAGAAC	640
OY	122	GluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgLIn	141
OY	142	TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln	161
Db	701	TGGCTGGAGTGGGGCGGTGAAGAATATGGCTTCCAGACGTCACATCTTGTATTCCAG	760
OY	162	AsnIleAspGlyLysGluLeuCysLysMethTyrLysAspAspPheGlnArgLeuThrPro	181
Db	761	AACATCGATGGGAAGAACTGTGCAGATGACCAAGAGACGACTTCCAGAGCTCACCCCC	820
OY	182	SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr	201
Db	821	AGTCAACAACGCGGACATCTTCTCCATCTCCACTCCACGACGTCAGACACTCTCTTCCA	880
OY	202	PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleGThrThrArgPro	221
Db	881	CATTTG-----ACTTCA-----GATATGTGTATAAGCTTACAAAACTCTCCA	925
OY	222	-----AspLeuProTyrGluGlnAlaIleArgSerAlaTrp	233
Db	926	CGGTTAATGATGCTAGAAACACAGATTATTACATATGAGCCGCCAGAGATCAGCCTGG	985
OY	234	ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal	252
Db	986	ACCGGTGACGGCCACCCACCCACCCAGTCGAAAGCTGTCTCAACCATCTCCCTCCACAGT	1045
OY	253	ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer	272
Db	1046	CCCAAAACTGAAAGACACGCTCTCAGTTAATCTTATTCAGATTCTTGGACCAACAAGT	1105
OY	273	SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeu	292
Db	1106	AGCGCGCTTGCMAATCCAGGACAGTGGCCAGATCCAGTCTTGGAGTTCTCCTGTGAGTCC	1165
OY	293	LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLys	312
Db	1166	CTGTGCGACACCTCCAACTCCAGCTGCATCCCTGGGGAAGCACCAACGGGGAGTTCAAG	1225
OY	313	MethTrpAspProAspGluValAlaIleArgTrpArgGluArgLysSerLysProAsnMet	332
Db	1226	ATGACGGATTCGACGAGAGGTGGCCGGCGCTGGGGAGAGCGGAAGCAAAACCCCAACATG	1285
OY	333	AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMethTyrLys	352
Db	1286	AACATGAGTAAAGCTCAGCGCGGCCCTCCGTTACTACTATGACAAAGACATCATGACCAAG	1345
OY	353	ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln	372
Db	1346	GTCATGAGGAGAGCGCTACGCTTACAAGTTGACATTCCACGGGATGCGCCAGCCCTCCAG	1405
OY	373	ProHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer	392
Db	1406	CCCCACCCCCCGGAGTCACTCTCTTACAAAGACCCCTCAGACCTCCCTCATATGGGCTCC	1465
OY	393	TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal	412
Db	1466	TATACACGCCCAACCAAGATGAACTTTTGGGCCCCCAGCCCTCCAGCCCTCCCGGTG	1525
OY	413	ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle	432
Db	1526	ACATCTTCCATTTTTTTGGTGGCCCAAAACCCATCTAGATTCACCAACATGGGGGTATA	1585
OY	433	TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	451

```

DB      1586  ACCCCCAACATAGAGCTCCACCACGACCATATGCTTCTCATCTGAGGCACTTACTAC 1642

RESULT 4
US-09-864-761-18410/c
; Sequence 18410, Application US/09864761
; Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weisheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecmca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263, 6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 18410
LENGTH: 567
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000163.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HEL1, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUATE 9.00e-99
OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUATE 0.00e+00
OTHER INFORMATION: NT HIT: M17254.1, EVALUATE 0.00e+00

```



```
Db 233 GCGTACAGTTCGACCTCCAGGATCGCCAGCGCCCTCCAGCCCGCCAGAGCA 174
Qy 379 SerMetYrlySyrProSerAspLeuProTyrMetSerTyrYHsAlaHsProIn 398
Db 173 TCTGTGACAGTACCCCTCCAGACCTCCGATACATGGGCTCTTATCACGCCACACAG 114
Qy 399 LysMetAspHeValAlaProHsProAlaLeuProValThrSerSerPhePhe 418
Db 113 AAGATGAACCTTGTGGGGCCCCCAGCCCTCCAGCCCTCCGATCATCTTCCAGTTT 54
Qy 419 AlaAlaProAsnProTyrTTPasSerProThrGlyIleTyrProAsn 435
Db 53 GCTGCCCAACCACTACTGGAATTCACCAACTGGGGGTATATACCCCAAC 3

RESULT 6
US-09-923-779-54
; Sequence 54, Application US/09923779
; Patent No. US20020076721A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.553
; CURRENT APPLICATION NUMBER: US/09/923, 779
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 356, 363, 417, 433, 442, 461, 463, 464, 469, 479, 485, 489,
; LOCATION: 537, 545
; OTHER INFORMATION: n = A,T,C or G
US-09-923-779-54

Alignment Scores:
Pred. NO.: 2,396-68 Length: 549
Score: 712.00 Matches: 128
Percent Similarity: 95.59% Conservative: 2
Best Local Similarity: 94.12% Mismatches: 5
Query Match: 29.13% Indels: 1
DB: 10 Gaps: 0

US-09-902-772-2 (1-451) x US-09-923-779-54 (1-549)
Qy 317 AspGluValAlaArgATGTPGlyGluArgLysSerLysProAsnMetAsnTyrAspLys 336
Db 2 GAGAGAGTGCGCCGCGCTGGGGAGAGCGAGAGCAACCCACATGACTAGATAG 61
Qy 337 LeuSerArgAlaLeuAlaGlyTyrTyrAspLysAsnIleMetThrLysValHsGlyLys 356
Db 62 CTCGCGCGCGCCCTCCCTTACTACTATGACAAAGACATCATGACCAAGTCCATGGGAG 121
Qy 357 ArgTyrAlaTyrLysPheAspPheHsGlyIleAlaGlnAlaLeuGlnProHsProPro 376
Db 122 CGCTACGCTTACAGATTCGACTTCACGGGATCGCCAGGCGCCCTCCAGCCCGCCG 181
Qy 377 GluSerMetYrlySyrProSerAspLeuProTyrMetSerTyrYHsAlaHs 396
Db 182 GAGTCATCTCTGTACAGTACCCCTCCAGACCTCCGATACATGGGCTCTATACAGCCAC 241
Qy 397 ProGlnLysMetAspHeValAlaProHsProAlaLeuProValThrSerSer 416
Db 242 CCACAGAGATGAACCTTGTGGGGCCCCCAGCCCTCCAGCCCTCCGATCATCTTCCAGT 301
Qy 417 PheHeAlaAlaProAsnProTyrTTPasSerProThrGlyIleTyrProAsnThr 436
Db 417 PheHeAlaAlaProAsnProTyrTTPasSerProThrGlyIleTyrProAsnThr 436
```

```
Db 302 TTTTGTGCTGCCCAACCACTGCAATTCACCAACTGGGGGTATATACCCNTCACT 361
Qy 437 ArgLeuProAlaAlaHsMetProSerHsLeu-GlyThrTyrTyr 451
Db 362 ANGCTCCCAACAGACCATATGCTTTTCATCTGTGGGGCACTTACTAC 407

RESULT 7
US-09-864-761-3705/c
; Sequence 3705, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180, 312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207, 456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632, 366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236, 359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234, 687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608, 408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774, 203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3705
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000021.2
; OTHER INFORMATION: EXPRESSED IN B474, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
US-09-864-761-3705
```

Alignment Scores:

Pred. No.:	9,466-62	Length:	420
Score:	650.00	Matches:	114
Percent Similarity:	98.31%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	2
Query Match:	26.60%	Indels:	0
DB:	10	Gaps:	0

US-09-902-772-2 (1-451) x US-09-864-761-3705 (1-420)

```

OY      334  TyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysVal 353
          |||
DB      418  TACGATAGCTCAGCGCGCCCTCCGTTACTACTATGACAGAACATCATGACCAAGGTC 359
          |||
OY      354  HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
          |||
DB      358  CATGGGAAGCGCTACGCTTACAGAGTTCCACTTCCACGGGAGTCGCCAGGCCCTCCAGGCC 299
          |||
OY      374  HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
          |||
DB      298  CACCCCGCGAGTCTCTCTGTGACAGTACCCTCAGACCTCCGTACATGGGCTCTAT 239
          |||
OY      394  HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
          |||
DB      238  CACGCCACCCACAGAAAGATGAACTTTGTGGGCGCCCACTCCAGCCCTCCCGTGACA 179
          |||
OY      414  SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyLysTyr 433
          |||
DB      178  TCTTCCAGTTTTTTTGGTGGCCCAAAACCCATACGTGAATTCACCAACTGGGGGTATATAC 119
          |||
OY      434  ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
          |||
DB      118  CCCAACACTAGGCTCCCCACCAAGCATATGCTCTCTCATCTGGGCACTTACTAC 65
          |||

```

RESULT 8

US-09-864-761-1653/c

Sequence 1653, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

```

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```

```

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 1653
LENGTH: 454
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000163.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95

```

US-09-864-761-1653

Alignment Scores:

Pred. No.:	1,066-61	Length:	454
Score:	650.00	Matches:	114
Percent Similarity:	98.31%	Conservative:	2
Best Local Similarity:	96.61%	Mismatches:	2
Query Match:	26.60%	Indels:	0
DB:	10	Gaps:	0

US-09-902-772-2 (1-451) x US-09-864-761-1653 (1-454)

```

OY      334  TyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysVal 353
          |||
DB      452  TACGATAGCTCAGCGCGCCCTCCGTTACTACTATGACAGAACATCATGACCAAGGTC 393
          |||
OY      354  HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
          |||
DB      392  CATGGGAAGCGCTACGCTTACAGAGTTCCACTTCCACGGGAGTCGCCAGGCC 333
          |||
OY      374  HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
          |||
DB      332  CACCCCGCGAGTCTCTGTGACAGTACCCTCAGACCTCCGTACATGGGCTCTAT 273
          |||
OY      394  HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
          |||
DB      272  CACGCCACCCACAGAAAGATGAACTTTGTGGGCGCCCACTCCAGCCCTCCCGTGACA 213
          |||
OY      414  SerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGlyLysTyr 433
          |||
DB      212  TCTTCCAGTTTTTTTGGTGGCCCAAAACCCATACGTGAATTCACCAACTGGGGGTATATAC 153
          |||
OY      434  ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
          |||
DB      152  CCCAACACTAGGCTCCCCACCAAGCATATGCTCTCTCATCTGGGCACTTACTAC 99
          |||

```

RESULT 9

US-09-918-995-23356

Sequence 23356, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED


```

: TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918, 995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235, 076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 38054
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 23356
: LENGTH: 467
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)...(467)
: OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23356

Alignment Scores:
Pred. No.: 2,24e-60 Length: 467
Score: 638.00 Matches: 120
Percent Similarity: 97.66% Conservative: 5
Best Local Similarity: 93.75% Mismatches: 3
Query Match: 26.10% Indels: 0
DB: 12 Gaps: 0

US-09-902-772-2 (1-451) x US-09-918-995-23356 (1-467)
QY 1 MetAlaSerThrIleuYsguAlaLeuSerValAlaSerGluAspGlnSerLeuPheGlu 20
Db 84 ATGCCCGACACTATTAAAGAGGCTTATCACTGTGTGATGAGGACCACTGCTTTGAG 143
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40
Db 144 TGTCTCTACGAGAGCCACACCTGCTAAGACAGATGACCGGCTCTCTCCAGGAGAC 203
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60
Db 204 TATGACAGACTTCCAGATGAGCCACGCGTCCCTCAGCAGATGCTGTCTCAACCC 263
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 264 CCACCCAGGCGTCCACCAATGATGATTAACCTTACGAGGATGCTCAAGAAC 323
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 324 TCTCTGATGATGACGTGTGGCCAAAGCGGGAAGATGTGGGCGACCCAGACCGGT 383
QY 101 GlyMetAspTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThr 120
Db 384 GGGATGACTACGCGACCTACATGAGAGAGACATGCCACCCCAACATGACCAAG 443
QY 121 AsnGluArgArgValIleValPro 128
Db 444 AACGAGCGCAGAGTATCTGTGCCA 467

RESULT 10
US-09-850-799-1
: Sequence 1, Application US/09850799
: Patent No. US20020090647A1
: GENERAL INFORMATION:
: APPLICANT: Demerits, Evan S.
: APPLICANT: Eyodoto, Dmitry V.
: APPLICANT: Hendricks, Timothy J.
: TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
: TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
: FILE REFERENCE: CASE-03828
: CURRENT APPLICATION NUMBER: US/09/850, 799
: CURRENT FILING DATE: 2001-05-08
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/360, 779
: PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-26
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0

```

```

: SEQ ID NO 1
: LENGTH: 1752
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (112)...(1131)
US-09-850-799-1

Alignment Scores:
Pred. No.: 1.06e-45 Length: 1752
Score: 511.00 Matches: 123
Percent Similarity: 50.69% Conservative: 23
Best Local Similarity: 42.71% Mismatches: 63
Query Match: 20.91% Indels: 79
DB: 10 Gaps: 10

US-09-902-772-2 (1-451) x US-09-850-799-1 (1-1752)
QY 209 ValTyrProGluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAla 228
Db 369 GTGCCCGCCCGTGGCGGTCGCCCGCCATCTCCACCGCCCA----- 407
QY 229 ArgArgSerAlaTyrThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSer 248
Db 408 -----GTCCCGCGCAGAGATGAGACAGCGCGCACCTCCAGCCCTGCT 452
QY 249 SerSerThrValProLysThrGlnAspGlnArg-----Pro 260
Db 453 GATCACATGATGACTTACCAGATCCCGTGGAGATGCTTTTAAAGAGGAGAACGCC 512
QY 261 GlnLeu-AspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlyse 280
Db 513 GAGCTGGGGGCG-----CTGAGCCCTGGCGGTACAGAAA-----GGCAG 551
QY 280 rGlyGlnIleGlnLeuThrGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSerAs 300
Db 552 CGGGCAGATCAGATGTGGAGATTCTACATGAGACTCTGTCGACAGCCGGAACCGCGG 611
QY 300 nCysIleThrProGluGlyThrAsnGlyLysPheLysMetThrAspProAspGluValAl 320
Db 612 CTGATCGCGTGGGAGGCGCGCCAGCGCGACTTCAAGTCAACGCCACCGACGCTGGC 671
QY 320 AArgArgTyrGlyLysLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAl 340
Db 672 GCGAGCGTGGGAGGAGCAAGAACAGCCCAATATGACTACGACAACTAGTGCAGC 721
QY 340 AleuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAlaTy 360
Db 732 ACTGCGTACTACTACACAAACAAATCATGACCAAGTGCACGCGCAAGCGTACGCTA 791
QY 360 rLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisPro----- 375
Db 792 CCGCTTGACTTCCAGAGGCGCTGGCAGACAGCTTGGCAGCCACCGCGCGACCCACGC 851
QY 376 -----ProGluSerSerMetTyrLysTyrPr 384
Db 852 CGCCGCTGCGCGCGCGCAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 911
QY 384 oSerAspLeu-----ProTyrMetSerSerTyrHisAlaHisProGlnLys 399
Db 912 GGCTGTGTGGCTGCACCTGCCCTTCCCGGCGCTCTCC-----AA 950
QY 399 sMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAl 419
Db 951 ACTCAACCTTATGACAGCTCGCGCGCGCGCGCGCGCGCGCGCTTCTCT----- 999
QY 419 AlaProAsnProTyrTyrAsnSerPro-----ThrGlyGln 431
Db 1000 -----TACTGGCGCTGTGCCAAGCGCACCGCGCGCGCGCGCGCGCGCTGC 1046
QY 431 yIleTyrProAsnThrArgLeuPro-----AlaAlaHisMetPr 444

```


Db	Qy	Db
1047	444	1107
GCCTTACCAACCCCGGGGCTTGGACGCCCTCCCGGGCCCTTTGGCGGGGTGGCGCCGCGC	0SerHisLeuGlyThrTyrTyr 451	
	:::	
	TTGCGACTTGGGGGTCATTAT 1128	

```

RESULT 11
US-09-954-531-955
; Sequence 955, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 955
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-955

```

Pred. No. :	2,14e-38	Length:	2269
Score:	445.50	Matches:	127
Percent Similarity:	40.96%	Conservative:	43
Best Local Similarity:	30.60%	Mismatches:	105
Query Match:	18.23%	Indels:	141
DB:	11	Gaps:	14

US-09-902-772-2 (1-451) x US-09-954-531-955 (1-2269)

```

QY      122 GUARGAGValIleValProAlaAspProThriLeuTPSerThrAspHisValArgIn 144
       :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      571 CAGCGGCGCCTGGCATTCGAAGAACCCTGGCTGTGCATGAGCAACAGATTGCCAG 630
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      142 TPRLeuGLUTPRalavalLysGlutryrGlyLeuProkAspvalAspIIleLeuPheGln 161
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      631 TGGCTTCTTGGGCCACCAATGACTGCTGTGGAGCGGAAATCGCACAGAGTTG-- 687
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      162 AsnIleaspgLyLysGIleucYslsMethTrirlysaspAspPhagInArgLeuthPro 181
       :::::  :::::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      688 GGCATGAATAGCCACAGATGCTGTGNAACTTGSCAAGAACGCTTTCGGAGCTGCACCT 747
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      182 SerTyrrAsnalAspIIleuleuSerHisIstueHistryIleuarGluarGglyalaThr 201
       :::::  :::::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      748 GACTTGTGGGATGCATTCTCTGGGAMATCTGGAGCAAGCAAGATCAAGAAAACCAA-- 804
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      202 PheilPheProAsnThrSerValtyzProGluIalThrGlnatrgIlethrThrArgPro 221
       :::::  :::::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      805 -----GAAAGACAGAAAGATCATATGAAAGAAAAATTCACACCTGCACCTCCGTTCT 855
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      222 -----AspleuProTyrr-GluGlnal 228
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db      856 CATGTGATTAAcGCAATATACATTAGTTTGGCAGACAGCAGCGGCCCTATNGAATCAG 915
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      228 aargyrg-----SeratITPrThSer-----HisSerHisProTh 240
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      916 ACACGAATTAACCCAAAGGCGGCTCTCTGGACAGCATGTGTCCGGCTCCACACCCAGC 975
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

[illegible]

```

; LENGTH: 2188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-300A-1716

Alignment Scores:
Pred. No.: 2,61e-38 Length: 2188
Score: 444.50 Matches: 125
Percent Similarity: 41.28% Conservative: 43
Best Local Similarity: 30.71% Mismatches: 105
Query Match: 18.19% Indels: 135
DB: 11 Gaps: 13

US-09-902-772-2 (1-451) x US-09-902-300A-1716 (1-2188)

QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
   ::::::::::: ::::: ::::: ::::: :::::
Db 491 CAGGGGGCCGCGGCGCATTCACAAAGAACCCCGGCGTGTGAGTGAAGCAACAGTATGCCAG 550

QY 142 TrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPheGln 161
   ||||| ||||| ||||| ||||| |||||
Db 551 TGGCTTCTGCTGGCCACCAATGATTCAGTCTGTGAACGTGAATCTGCAGAGGTTTC--- 607

QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
   ::::::::::: ||||| ::::: ||||| ::::: |||||
Db 608 GGCATGAAATGGCCAGATGCTGTGTACCTTGGCAAGGACGCTTCTGAGCTGGCACCT 667

QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
   ::::: ||||| ||||| ::::: :::::
Db 668 GACTTTGTGGTGATCTCTCTGAGAACATCTGGAAACATGATCAAAAGAAACCAA--- 724

QY 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrArgPro 221
   ::::: ||||| ||||| ::::: |||||
Db 725 -----GAAAAGACAGAAAGATCATATGAAAGAAAATTCACACCTCACCTCGTCTCT 775

QY 222 -----AspLeuProTyr-GluGlnAl 228
   ::::: ||||| ||||| ::::: |||||
Db 776 CATTTGGATTAAACAGCAATACATTAGTTTGGCACAGACGCGCCCTATGGAATGCAG 835

QY 228 AlaArgArg-----SerAlaTrpThrSer-----HisSerHisProth 240
   ||||| ::::: ||||| |||||
Db 836 ACACAGAAATTAACCCCAAGGCGGCTCTGAGACAGATGTGCGGCTCCACACCCAGC 895

QY 240 RginSerLysAlaThrGln-----ProSerSe 249
   ||||| ::::: ||||| |||||
Db 896 GTACTCAGCTCTGAGCAGGAGTTTCAGATGTTCCCAAGTCTCGGCTCAGCTCCGTACG 955

QY 249 rSerThr----- 251
   |||||
Db 956 GTCACTACTGCTCTGTCACTGACGACTCCAGGACAGCAACTTGAATTGGCTCACCAC 1015

QY 252 -----ValProLysThr-----GluAs 257
   ::::: ||||| |||||
Db 1016 AATTCTGGAGCTCCCAAGACACAGACTCCCTGAGAGAGGTTGGGACAGCTTCGAGAGC 1075

QY 257 pGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro----- 270
   ||||| ||||| |||||
Db 1076 TCAGACTCCCTCTCCAGTCTGGAACAGCAGTCTGCTTGATGATGCAACGGGTT 1135

QY 270 ----- 270
   |||||
Db 1136 CCTTCCCTTCAGAGCTTCGAAGATGACTGACGCCAGCTCTCTGCTCAATTAAGCAACC 1195

QY 271 -----ThreSerSerArgLeuAlaAsnPro----- 278
   ||||| ||||| |||||
Db 1196 ATGCTTTTCAAGATTACATCAAGAGAGG-AGTGAACCCAGTGGAGCAAGCAACACAGT 1254

QY 279 -----GlySerGlyGlnIleGlnLeuTrpGln 288
   ||||| ||||| ||||| |||||
Db 1255 TATACCTGCAGCTGTGTGGCCGCTTCACAGGAAGAGTGAATTCAGCTGTGGCAGTT 1314

QY 288 eLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpIleGluGlyThrAs 308
   ||||| ||||| ||||| ||||| |||||

```

```

Db 1315 TCTCTGGAGCTGCTATCAGACAAATCTCCAGTCAATTCATCAGCTGGAGTGCAGCG 1374
QY 308 nGlyLuphElyMetThrAspProAspGluValAlaArgArgTrpGlyArgLysSe 328
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1375 ATGGAGTTTAAGTCTCCGCCGACCCCGATGAGTGGCCGCGGTGGGAAAGAGAAATA 1434

QY 328 rLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAs 348
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1435 TAAGCCCAAGATGAAGTACAGAGAGTACAGCCGGGGCTTACGCTACATTCAGCAAGAA 1494

QY 348 nIleMetThrLysValHisGlyLysArgTyrAlaTyrLysPhe----- 362
   ||||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1495 CATATCCACAAGACGTGCGGGAAGCGCTACGTGACCGCTCTGTGCGACCTCCAGAA 1554

QY 363 -----AspPheHisGlyIleAlaGlnAlaLeuGlnProHisPr 375
   ::::: ||||| ||||| ||||| |||||
Db 1555 CTTCGTGGGTTTCACGCCCGGAGAACTGCAGCCATCTCTG---GGCTTCAGCCCGAAC 1611

QY 375 oProGluSerSerMetTyrLysTyrProSerAspLeuPro---TyrMetSerSerTyrH 394
   ::::: ||||| ||||| ||||| |||||
Db 1612 GGAGGACTGAGGTGCGCGGACACACCTGAGCCGCCAGGCTGCGAGTGAAGTGGGA 1671

QY 394 sAlaHisProGlnLysMet 400
   ||||| ||||| ||||| |||||
Db 1672 AGCCATCTGCACCAAGCTG 1690

RESULT 13
US-10-033-528-1716
; Sequence 1716, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSTS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ. ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 1716
; LENGTH: 2188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-1716

Alignment Scores:
Pred. No.: 2,61e-38 Length: 2188
Score: 444.50 Matches: 125
Percent Similarity: 41.28% Conservative: 43
Best Local Similarity: 30.71% Mismatches: 105
Query Match: 18.19% Indels: 135
DB: 15 Gaps: 13

US-09-902-772-2 (1-451) x US-10-033-528-1716 (1-2188)

QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
   ::::::::::: ::::: ::::: ::::: :::::
Db 491 CAGGGGGCCGCGGCGCATTCACAAAGAACCCCGGCGTGTGAGTGAAGCAACAGTATGCCAG 550

QY 142 TrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPheGln 161
   ||||| ||||| ||||| ||||| |||||
Db 551 TGGCTTCTGCGGCGCACCAATGATTCAGTCTGTGAACGTGAATCTGCAGAGGTTTC--- 607

QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
   ::::::::::: ||||| ::::: ||||| ::::: |||||
Db 608 GGCATGAAATGGCCAGATGCTGTGTACCTTGGCAAGGACGCTTCTGAGCTGGCACCT 667

QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
   ::::: ||||| ||||| ::::: :::::
Db 668 GACTTTGTGGTGATCTCTCTGAGAACATCTGGAAACATGATCAAAAGAAACCAA--- 724

```

```

QY 202 PheIleheProAsnThrSerValTyProGluValaThrGlnArgIleThrThrPro 221
Db 725 -----GAAAGACAGAAAGATCATATGAAACAAAATTCACACCTCAGCTCGTCT 775
QY 222 -----AspLeuProTyR-GluGlnAl 228
Db 776 CATTGATTAACACATACATATAGTTTGGACAGAGAGCGCCCTATGAAATGCAG 835
QY 228 aArgArg-----SerAlaTrpThrSer-----HisSerHisProth 240
Db 836 ACACAGATTATACCCAAAGCGGCTCTGAGACAGATGTGTCCGCTCCACACCCAGC 895
QY 240 rGlnSerLysAlaThrGln-----ProSerSe 249
Db 896 GTACTCAGCTCTGAGCAGAGATTTCAGATGTTCACCAAGTTCGGCTCAGCTCCGTAC 955
QY 249 rSerThr----- 251
Db 956 GTCACTACTACTGCTCTGTCACTCAGTCCAGCAGCAACTGAAATTTGCTCACCAC 1015
QY 252 -----ValProLysThr-----GluAs 257
Db 1016 AATTCTGGAGCTCCCAAGACACGACTCCCTGAGAACGCGTGGACAGCTTCGAGAGC 1075
QY 257 pGlnArgProGlnLeuAspProTyRglnIleLeuGlyPro----- 270
Db 1076 TCAGACTCCCTCCCTCCAGTCTCTGAGAACGACGATGCTGTGATGTGCAACGGGTT 1135
QY 270 ----- 270
Db 1136 CCTTCCTTCAGAGACTTCGAAGATGACTGAGCAGCACTCTCTCTCATTAAGCAAC 1195
QY 271 -----ThrSerSerArgLeuAlaAspPro----- 278
Db 1196 ATGCTTTTCAAGATTATCATCCAGAGAG-AGTGACCAGGTGAGCAGCAACCAAGT 1254
QY 279 -----GlySerGlyGlnIleGlnLeuTrpGlnPh 288
Db 1255 TATACCTGACAGCTGTGCTGGCCGCTTCACAGAGAGTGGACGCTTACGCTGGCAGTT 1314
QY 288 eLeuLeuGlnLeuSerAspSerAsnSerAsnLysIleThrTrpGlnLysThrAs 308
Db 1315 TCCTCTGAGCTGCTATCAGCAAAATCTCCAGTATCATCATGCTGAGCTGGAGAGCG 1374
QY 308 nGlyGlnPheLysMetThrAspProAspGluValAlaArgArgTrpGlnLysSe 328
Db 1375 ATGGAGATTAAAGCTGCGCGAGCCGATGAGGTGGCCGCGGGGAAAGAGAAAA 1434
QY 328 rLysProAsnMetAsnTyRAspLysLeuSerArgAlaLeuArgTyRTrpLysAspLysAs 348
Db 1435 TAAOCCCAAGATGACACTACAGAGAGCTGAGCCGGGCTTACGCTACTATTACGACAA 1494
QY 348 nIleMetThrLysValHisGlyLysArgTyRAlaTyRAspHe----- 362
Db 1495 CACTATCCAAAGACGTCGGGGAAGCGCTACGCTCCGCTGCTGGCAGCTCCAGAA 1554
QY 363 -----AspPheHisGlyIleAlaGlnAlaLeuGlnProHisPr 375
Db 1555 CTTCGCTGGGCTTCACGCCGAGAACTGACAGCCCATCTCTG--GGCTCCAGCCGACAC 1611
QY 375 oProGlnSerSerMetTyRTrpProSerAspLeuPro-----TyrMetSerSerTyRnI 394
Db 1612 GAGAGACTGAGAGTCGCCGGAGACCACTTGAACGCCGCCAGGCTGTGAGTGTGGA 1671
QY 394 sAlaHisProGlnLysMet 400
Db 1672 AGCCCATCTGACCACTG 1690

```

RESULT 14
 US-09-920-300A-1693
 ; Sequence 1693, Application US/09920300A
 ; Patent No. US20020136728A1

```

; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920.300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1693
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1693

Alignment Scores:
Pred. No.: 2,74e-38 Length: 2268
Score: 444.50 Matches: 125
Percent Similarity: 41.28% Conservative: 43
Best Local Similarity: 30.71% Mismatches: 105
Query Match: 18.19% Indels: 135
DB: 11 Gaps: 13

US-09-902-772-2 (1-451) x US-09-920-300A-1693 (1-2268)
QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 571 CAGCGGCGCTGGGCAATTCACAAAGAACCCCTGCTGTGAGTGAACAGTGAAGTGCAG 630
QY 142 TrpLeuGlnTrpAlaValLysGlnTyRglnLeuProAspValAlaPheLeuPheGln 161
Db 631 TGCTTCTCTGGGCGACCAATGAGTTCAGTGTGTGTAACGTAATTCGACAGAGGTTG 687
QY 162 AsnIleAspGlyLysGlnLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 688 GGCATGAAATGGCCGATGCTGTGTATACCTTGGCAGAGACGCTTCTGGAGCTGGCAGCT 747
QY 182 SerTyRAsnAlaAspIleLeuLeuSerHisLeuHisTyRLeuArgGlnArgGlyAlaThr 201
Db 748 GACTTGTGGGTGACATTCCTCTGGAGAACATCTGGAGCAATGATCAAAAGAAACCA 804
QY 202 PheIleheProAsnThrSerValTyProGluValaThrGlnArgIleThrThrArgPro 221
Db 805 -----GAAAGACAGAAAGATCATATGAAACAAAATTCACACCTCAGCTCGTCT 855
QY 222 -----AspLeuProTyR-GluGlnAl 228
Db 856 CATTGATTAACACATACATATAGTTTGGCAGAGAGAGCGCCCTATGAAATGCAG 915
QY 228 aArgArg-----SerAlaTrpThrSer-----HisSerHisProth 240
Db 916 ACACAGATTATACCCAAAGCGGCTCTGAGACAGATGTGTCCGCTCCACACCCAGC 975
QY 240 rGlnSerLysAlaThrGln-----ProSerSe 249
Db 976 GTACTCAGCTCTGAGCAGAGATTTCAGATGTTCACCAAGTTCGGCTCAGCTCCGTAC 1035
QY 249 rSerThr----- 251
Db 1036 GTCACTACTACTGCTCTGTCACTCAGTCCAGCAGCAACTGAAATTTGCTCACCAC 1095
QY 252 -----ValProLysThr-----GluAs 257
Db 1096 AATTCTGGAGCTCCCAAGACACGACTCCCTGAGAACGCGTGGACAGCTTCGAGAGC 1155
QY 257 pGlnArgProGlnLeuAspProTyRglnIleLeuGlyPro----- 270
Db 1156 TCAGACTCCCTCCCTCAGTCTCTGAGAACGACGATGCTGTGATGTGCAACGGGTT 1215
QY 270 ----- 270

```


Oy	375	oproglnsersermetylrlystyrproserarplpPro---tyrmetlsersertyni	394
	:::	:::	:::
Db	1692	ggaagactgtaggtgtccgggacaccccttagaacggccggccagagctctgtagatgtagtgga	1751
Oy	394	salahisproglinsymet	400
		:::	
Db	1752	agcccatctgtaccacgctg	1770

Search completed: July 28, 2003, 10:09:32
Job time : 297.484 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:20:09 : Search time 2979.8 Seconds

(Without alignments)
3805.372 Million cell updates/sec

Title: US-09-902-772-2

Perfect score: 2444

Sequence: 1 MASTIKKALSVSEDSLEFE.....IYPNTRLPAAHMPSHLCTYY 451

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+D2n.model -DEV=xlh
-Q=/cgn2_1/USPFO.spool/US09902772/runat_23072003_093658_14926/app_query.fasta.1.1294
-DB=Pending_Patents_NA_Main -QEXT=fastap -SUFFIX=rimpm -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -LOCAL=IGM-200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NOM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09902772.GCGN.1.1.4066.6runat.23072003.093658.14926 -NCPU=6 -ICPU=3
-NO_MAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main.*

1: /cgn2_6/pdata/1/pna/US097A.COMB.seq.*
2: /cgn2_6/pdata/1/pna/US06.COMB.seq.*
3: /cgn2_6/pdata/1/pna/US07.COMB.seq.*
4: /cgn2_6/pdata/1/pna/US08.COMB.seq.*
5: /cgn2_6/pdata/1/pna/US081.COMB.seq.*
6: /cgn2_6/pdata/1/pna/US082.COMB.seq.*
7: /cgn2_6/pdata/1/pna/US083.COMB.seq.*
8: /cgn2_6/pdata/1/pna/US084.COMB.seq.*
9: /cgn2_6/pdata/1/pna/US085.COMB.seq.*
10: /cgn2_6/pdata/1/pna/US086.COMB.seq.*
11: /cgn2_6/pdata/1/pna/US087.COMB.seq.*
12: /cgn2_6/pdata/1/pna/US088.COMB.seq.*
13: /cgn2_6/pdata/1/pna/US089.COMB.seq.*
14: /cgn2_6/pdata/1/pna/US090.COMB.seq.*
15: /cgn2_6/pdata/1/pna/US091.COMB.seq.*
16: /cgn2_6/pdata/1/pna/US092.COMB.seq.*
17: /cgn2_6/pdata/1/pna/US093.COMB.seq.*
18: /cgn2_6/pdata/1/pna/US094.COMB.seq.*
19: /cgn2_6/pdata/1/pna/US095A.COMB.seq.*
20: /cgn2_6/pdata/1/pna/US095B.COMB.seq.*
21: /cgn2_6/pdata/1/pna/US095C.COMB.seq.*
22: /cgn2_6/pdata/1/pna/US096A.COMB.seq.*
23: /cgn2_6/pdata/1/pna/US096B.COMB.seq.*
24: /cgn2_6/pdata/1/pna/US096C.COMB.seq.*
25: /cgn2_6/pdata/1/pna/US096D.COMB.seq.*
26: /cgn2_6/pdata/1/pna/US096E.COMB.seq.*
27: /cgn2_6/pdata/1/pna/US096F.COMB.seq.*
28: /cgn2_6/pdata/1/pna/US097A.COMB.seq.*

Result No.	Score	Query Match	Length	ID	Description
1	2444	100.0	1447	US-09-902-772-1	Sequence 1, App11
2	2414.5	98.8	1528	US-09-902-772-3	Sequence 3, App11
3	2197	89.9	3309	US-09-652-109-9617	Sequence 9617, Ap
4	2197	89.9	28	US-09-716-953-1936	Sequence 1936, Ap
5	2197	89.9	3309	US-09-721-589-5458	Sequence 5458, Ap
6	2197	89.9	29	US-09-726-806-4625	Sequence 4625, Ap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```

7 2197 89.9 3309 29 US-09-726-811-4920 Sequence 4920, Ap
8 2182 89.3 1389 35 US-09-948-941-233 Sequence 233, App
9 2182 89.3 3166 1 PCT-US02-23913-98 Sequence 98, Appl
10 2182 89.3 3166 30 US-09-784-356-26 Sequence 26, Appl
11 2182 89.3 3166 38 US-10-021-660-26 Sequence 26, Appl
12 2182 89.3 3166 42 US-10-205-823-98 Sequence 98, Appl
13 1991 81.5 2172 80 US-60-360-207-355 Sequence 355, App
14 1810 74.1 3126 18 US-09-442-384A-475 Sequence 475, App
15 1810 74.1 3126 18 US-09-442-384B-475 Sequence 475, App
16 1624.5 66.5 1359 39 US-10-035-832-1210 Sequence 1210, Ap
17 1624.5 66.5 1359 39 US-10-052-482-198 Sequence 198, App
18 1624.5 66.5 2938 7 US-08-343-443-9 Sequence 3, Appl1
19 1624.5 66.5 2957 1 PCT-US02-18947-710 Sequence 710, App
20 1624.5 66.5 2957 38 US-10-007-926A-295 Sequence 295, App
21 1624.5 66.5 2957 38 US-10-035-832-1209 Sequence 1209, App
22 1624.5 66.5 2957 39 US-10-052-482-197 Sequence 197, App
23 1624.5 66.5 2957 41 US-10-172-118-710 Sequence 710, App
24 1624.5 66.5 3298 25 US-09-652-814-9467 Sequence 9467, App
25 1624.5 66.5 3298 27 US-09-699-999-6013 Sequence 6013, App
26 1624.5 66.5 3298 28 US-09-716-953-1954 Sequence 1954, App
27 1624.5 66.5 3298 28 US-09-716-990-1139 Sequence 1139, App
28 1624.5 66.5 3298 29 US-09-721-589-5554 Sequence 5554, App
29 1624.5 66.5 3298 29 US-09-726-171-2148 Sequence 2148, App
30 1624.5 66.5 3298 29 US-09-726-806-4681 Sequence 4681, App
31 1624.5 66.5 3372 17 US-09-396-970-7615 Sequence 7615, App
32 1613.5 66.0 1359 38 US-10-035-832-1207 Sequence 1207, App
33 1613.5 66.0 1729 14 US-09-053-832-1206 Sequence 1206, App
34 1613.5 66.0 1729 38 US-10-035-832-1206 Sequence 1206, App
35 1613.5 66.0 1729 38 US-10-052-482-194 Sequence 194, App
36 1613.5 66.0 1729 39 US-10-052-482-194 Sequence 194, App
37 1613.5 66.0 1729 80 US-60-360-207-10845 Sequence 10845, A
38 1611.5 65.9 3110 76 US-60-324-185-33223 Sequence 33223, A
39 1595 65.3 1307 61 US-60-172-373-5267 Sequence 5267, App
40 1536.5 62.9 1932 14 PCT-US02-25766-2596 Sequence 2596, App
41 1536.5 62.9 1932 14 US-09-053-3758-327 Sequence 327, App
42 1536.5 62.9 1932 18 US-09-442-384A-530 Sequence 530, App
43 1536.5 62.9 1932 18 US-09-442-384B-530 Sequence 530, App
44 999 40.9 661 29 US-09-726-806-3434 Sequence 3434, App
45 965 39.5 1431 1 PCT-US02-06518-147 Sequence 147, App

```

ALIGNMENTS

```

RESULT 1
US-09-902-772-1
: Sequence 1, Application US/09902772
: GENERAL INFORMATION:
: APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: FILE REFERENCE: Chugai Selyaku Kabushiki Kaisha 5001
: CURRENT APPLICATION NUMBER: US/09/902,772
: CURRENT FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/08/878,177
: PRIOR FILING DATE: 1997-06-18
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1447
: TYPE: DNA
: ORGANISM: C-11 gene, c-erg gene w/ deletion, chicken DNA
US-09-902-772-1

```

Alignment Scores:

```

Pred. No.: 2,06e-181 Length: 1447
Score: 2444.00 Matches: 451
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 34 Gaps: 0

```

US-09-902-772-2 (1-451) x US-09-902-772-1 (1-1447)

```

Qy 1 MetalaserThrilleygluaAlaLeuSerValSerGlnsPginSerLeupheGlu 20
Db 63 ATGCACAGCACTATTAAAGCAATATATAGTGTGATGAGTAAGCAAGCTCTGTTGAG 122
Qy 21 CysAlaIyrlGlySerProHlsLeuAlaIySrlThGlnMetThrAlaSerSerSerGlu 40
Db 123 TGTCCACAGATGAGTCCCAACCTTGCAAAACAGAAATGACAGCTCTTCCAGTAA 182
Qy 41 TyrcIleGlnThrSerLysMetSerProArgValProGlnGlnsPrrPleuSerGlnPro 60
Db 183 TATGGCAAAACATTAAGATGATGAGCCCGCGCTTCCACAGAGACTGTTATCAGCCC 242
Qy 61 ProAlaArgValThrIleLysMetGluCysAnProAsnGlnValAsnGlySerArgAsn 80
Db 243 CCGGCACAGATTACATTAAGATGAGTGAATGTAACCAACAGGTTAATGAGTGAAGAAAT 302
Qy 81 SerProAspAspCysSerValAlaIyGlyGlyMetValSerSerSerAspVal 100
Db 303 TCACCTGATGACTCCACCGTGGCAAAAGAGGAAATGTTACACTTACAGACAAATGTT 362
Qy 101 GlyMetAsnTyrglySerTyrmGluGluIshIstleProProProAsnMetThr 120
Db 363 GGGATGACCTATGAGACTCATGAGAGAGAGATATTCGCTCCAAATATACAAAC 422
Qy 121 AsnGluArgValIleValProAlaAspProThrLeuThrAspHisValArg 140
Db 423 AATGACGAAGAGATTATGTCGACAGATCTCTGTTATGAGACAGACAGATGACGG 482
Qy 141 GlnTrpLeuGlnTrpAlaValIySrlGlyLeuProAspValAspIleLeuPhe 160
Db 483 CAGGGGTGGAGTGGGAGTGAAGAGATAGTCTTCACAGCGGACATCTGTTGTTTC 542
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 543 CAGACACTTATGAGGAAGAGCTGTGAATATGCCAAAGATGACTCCAGAGACTACG 602
Qy 181 ProSerTyrrAsnAlaAspIleLeuLeuSerHisLeuHisIstYrLeuArgIyAla 200
Db 603 CCGAGCTATTAAGCAGATATCTCTGTCACACTTACATCACTCAGAGAGAGAGGCC 662
Qy 201 ThrPheIlePheProAsnThrSerValTyrrProGlnAlaThrGlnArgIleThrArg 220
Db 663 ACTTTATTTTCCAAATATACATCAGTTTACCACCAACAGCAACAAATTAACAAAGG 722
Qy 221 ProAspLeuProTyrgluGlnAlaIyArgSerAlaThrPrrHisSerHisSerProThr 240
Db 723 CCAATTTACTTATATGACCAAGCAGAGATCAGCGTACAGATCAGACACTCCACT 782
Qy 241 GlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlnAspGlnArgPro 260
Db 783 CAGTCAAAACACTACCAACCATCATCTTCAACAGTGCCTCAAAACAGAAACAGCCGCT 842
Qy 261 GlnLeuAspProTyrglnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
Db 843 CAGTTAATCTTATATCAGATCTTGGACCAACAGCCGCTTGGAAATCCAGGGAGT 902
Qy 281 GlyGlnIleGlnLeuTrpGlnIlePheLeuGlnIleuLeuSerAspSerSerAsn 300
Db 903 GGGCAGATACAGCATATGAGCATCTCTACTGAGAGCTTGTGCGACACTCCAACTCAAC 962
Qy 301 CysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThrAspProAspGluValAla 320
Db 963 TGCATCACCTGGGAGGCGCAAAATGGGAGTCAAGATGACAGACCTGTAAGTGGCT 1022
Qy 321 ArgArgTrpGlnGluAlaGlySerLysProAsnMetAsnTyrrAspLysLeuSerArgAla 340
Db 1023 CCGCGTGGGAGAGAGAAACCAACCTTAACATGAATATGACAAACCTACCCGCGCA 1082
Qy 341 LeuArgTrpTyrrTyrrAspLysAsnIleMetThrLysValHisGlyLysArgTyrrAlaTyrr 360
Db 1083 CTTCGCTACTACTATGACAAATATTTATGACTTAAGTATGATGAACCTATGCTTAC 1142
Qy 361 LysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlnSerSerMet 380

```



```

Db      1143 AAATGATTTCCAGGAATGGCTCAGGCCCTCAGGCTCAGCCTCAGATCATCATG 1202
Qy      381 TyrlstfyrProserAspLeuProTyrMetSerSerrYrHsAlahisProGlnLysMet 400
Db      1203 TACAAATACCATCAGACCTCCCTACATGATGTTCCATCCATGACACCCCAAGATG 1262
Qy      401 AsnheValAlaProHisProProAlaLeuProValThrSerSerSerpheAlaAla 420
Db      1263 AACCTTGTAGCTCCCATCCCTCCCTGCTTGGCCCTAACCTCATCAGCTTTTGGCTGCC 1322
Qy      421 ProasnProTyrTrpAsnSerProThrGlyLyleTyrProasnThrArgLeuProAla 440
Db      1323 CCTAATCCATCTAGGAATTCACCAATGAGAGCATCTACCCCATACAGGCTCCAGCT 1382
Qy      441 AlahisMetProSerHisLeuGlyThrTyrTyr 451
Db      1383 GCTCATATGCCCTTCCATCTTGGCACCTACTAC 1415

```

RESULT 2

```

US-09-902-772-3
: Sequence 3, Application US/09902772
: GENERAL INFORMATION:
: APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and genes of
: FILE OF INVENTION: the Proteins
: FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
: CURRENT APPLICATION NUMBER: US/09/902,772
: CURRENT FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/08/878,177
: PRIOR FILING DATE: 1997-06-18
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1528
: TYPE: DNA
: ORGANISM: c-erb gene, chicken DNA
US-09-902-772-3

```

Alignment Scores:

```

Pred. No.: 4.53e-179 Length: 1528
Score: 2414.50 Matches: 450
Percent Similarity: 94.14% Conservative: 0
Best Local Similarity: 94.14% Mismatches: 1
Query Match: 98.79% Indels: 27
DB: 34 Gaps: 1

```

US-09-902-772-2 (1-451) x US-09-902-772-3 (1-1528)

```

Qy      1 MetIasertHrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGlu 20
Db      63 ATGGCAAGCACTATTAAAGAAAGCATTTTCAGTGGTGAAGAACCCAGCTCTTTTGAG 122
Qy      21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerglu 40
Db      123 TGTGCTACGAGATGCGCCCACTTGCAGAAAGCAAGAAATGACAGCTCTCTCCAGTGA 182
Qy      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
Db      183 TATGGGCAACATCTCAAGATGAGCCCGCGCTTCCCGACAGCAGCTGTTATCAACAGCC 242
Qy      61 ProAlaIrrValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db      243 CCGGCCAGAGTTACATTAAAGATGAGGTGAACCAACAGGTTAATGGGTCAAGGAAT 302
Qy      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db      303 TCACCTGATGCTGCGACGCTGGCAAAAGAGGAAATGTTAGCAATTCAGACAATGTT 362
Qy      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120
Db      363 GGGATGAACATATGGAACCTACATGAGAGAGAGATATTCGCGCTCCAAATATGACAAAC 422

```

```

Qy      121 AsnGluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArg 140
Db      423 AATGACGAAGAGATTATGTGCCAGAGATCTTACGTATGAGACACAGACATGTACGG 482
Qy      141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLysPhe 160
Db      483 CAGTGGCTGAGTGGCGCAGTAAGAGATATGTTCTCCAGACGTCGACATCTTGTGTTCC 542
Qy      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db      543 CAGAACATTAATGGAAGAGATTGTGTAATGACCAAAAGATGACTTCCAGAGACTCAGC 602
Qy      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg 198.
Db      603 CCGAGCTATACGAGATATTCCTCTGTGACACCTTACACTATCCTCAGAGACATCCCTT 662
Qy      198 -----
Db      663 CCACATTTGACTTCAGATGATGTTGATTAAGGCTTACAAAACCTCCACGGTTAATGCAT 722
Qy      199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213
Db      723 GCTAGAAACACAGGAGAGCCACTTTATTTTCCAAATCATCATGTTTACCCAGAAACA 782
Qy      214 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233
Db      783 ACGAAAGAAATACACACAGGCCAGATTACCTTATATGACAGAGAGAGATCAGCTGG 842
Qy      234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253
Db      843 ACGAGTCACAGCCATCCCACTCAGCTCAAAAGCTCAACCAACATCATCTTCAACAGTCCC 902
Qy      254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
Db      903 AAACAGAACACAGCCCTCTCATGTTAGATCTTATACATCTTGGACGACACAGCAGC 962
Qy      274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 293
Db      963 GGTCTTCCAAATCCAGGAGAGTGGGCAATACAGTATGCGAGTTCCTACTGAGACTCTTG 1022
Qy      294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMet 313
Db      1023 TCGGACAGCTCCCAACTCAACTGATCATCTGGGAGGCGCAAAATGGGGATTCAMAATG 1082
Qy      314 ThrAspProAspGluValAlaArgArgTyrGlyLysLysSerLysProAsnMetAsn 333
Db      1083 ACAGACCTGATGAAGTGGCTGGCGGTGGGAGAGAGAAACCAACCTAACATGATGAC 1142
Qy      334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
Db      1143 TATGACAAACTCAGCCGCTGCACCTTCTACTACTATGACAAAATATTATGACTAAAGTT 1202
Qy      354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
Db      1203 CATGTATACAGCTATGCTCAAAATTTGATTTCCACGAAATGCGTCAAGCCCTCCAGCT 1262
Qy      374 HisProGlnGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
Db      1263 CACCTTCCAGAAATCATCATGTAACCAATACCAACCTCCCTCCACAGAGATTTCTTAC 1322
Qy      394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
Db      1323 CATGACACCCCAAGAGATGAACTTTGTAGCTGCCATCCCTCCCTGCTTGGCCGATAC 1382
Qy      414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 433
Db      1383 TCATCCAGCTTTTGTGTCCTTAATCATACATGGAATTCACCAACTGAGAGGATCTAC 1442
Qy      434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db      1443 CCCAAATACAGGCTGCGACGCTGCTCATATGCTTCCATCTTGGCACCTACTAC 1496

```

RESULT 3

US-09-652-109-9617
 : Sequence 9617, Application US/09652109
 : GENERAL INFORMATION:
 : APPLICANT: McCarthy, Sean A.
 : TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 : TITLE OF INVENTION: THEREFOR
 : FILE REFERENCE: 1600.1180-001
 : CURRENT APPLICATION NUMBER: US/09/652.109
 : CURRENT FILING DATE: 2000-08-30
 : PRIOR APPLICATION NUMBER: 60/151,128
 : PRIOR FILING DATE: 1999-08-30
 : NUMBER OF SEQ ID NOS: 10105
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 9617
 : LENGTH: 3309
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-652-109-9617

Alignment Scores:

Pred. No.:	1,31e-161	Length:	3309
Score:	2197.00	Matches:	410
Percent Similarity:	91.96%	Conservative:	13
Best Local Similarity:	89.13%	Mismatches:	23
Query Match:	89.89%	Indels:	14
DB:	25	Gaps:	4

US-09-902-772-2 (1-451) x US-09-652-109-9617 (1-3309)

QY 1 MetValSerThrIleuysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 128 ATGGCCGACACTATTAAAGAAAGCTTATCAGTTGTGATGAGGACCAAGTGTGTTGAG 187
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 188 TGTGCTACGGAAGCAGCAGCACTGCTAAGACAGATGACCGGCTCTCTCCAGGAC 247
 QY 41 TtyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 DB 248 TATGACAGACTTCCAGATGAGCCACGCGCTCCCTGACGAGATGCTGCTCAACCC 307
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProHisGlnValAsnGlySerArgAsn 80
 DB 308 CCACCCAGGGTCACCATCAAAATGAAATGTAACCTGACCCAGGTAATGCTCAAGGAC 367
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 368 TCTCTGATGATGACATGCTGCGCAAGGGGGAAGATGGTGGGACCCAGACCGCTT 427
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThr 120
 DB 428 GGGATGACACTACGCGACCTACATGAGAGAGACATGCGCACCCCAACATACACG 487
 QY 121 AsnGluArgValIleValAlaProAlaAspProThrLeuTrpSerThrAspHisValArg 140
 DB 488 AAGGAGCGCAGAGTATTCGTCGCCAGCATCTCTACGTATGAGATACAGACCATGCGG 547
 QY 141 GlnTrpLeuGluTrpAlaValLysGlyLysGlyLeuProAspValAlaPheLeuPhe 160
 DB 548 CAGTGGCTGAGTGGGCGGTGAAGAAATATGCGCTTCCACACGTCATCTTGTATTC 607
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 608 CAGAACATCGATGGGAGGAAGCTGTGCAAGATGACCAAGAGCGACTTCCAGAGGCTCAC 667
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAla 200
 DB 668 CCCAGCTACAAAGCCGACATCTTCTCTCAGATTCACACTACCTCAAGAGACCTCTT 727
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArg 220
 DB 728 CCACATTTG-----ACTTCA-----GATGATGTGTGATAAAGCCTTACAAACTCT 772

QY 221 Pro-----AspLeuProTyrGluGlnAlaArgSerAla 232
 DB 773 CCAGGTTAATGATGCTAGAAACACAGATTATACATATGACCCCGCAGAGATGACGCC 832
 QY 233 TyrPheSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThr 251
 DB 833 TGGACCGGTACAGCGCCACCCACGCGCCACATCGAAGCTCTCAACATCTCTCTCCACA 892
 QY 252 ValProLysThrGluAspGlnArgProGluLeuAspProTyrGlnIleLeuGlyProThr 271
 DB 893 GTGCCCAAACTGAAGACAGCGCTCCATGATGATCTTATCAGATTCCTGGACACACA 952
 QY 272 SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlu 291
 DB 953 AGTAGCGGCTTGCAATTCAGGACAGCGACATCCAGCTTGGCGAGTCTCTCGAGAG 1012
 QY 292 LeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGluThrAsnGlyLuphe 311
 DB 1013 CTCTGTGAGACGCTTCAACTCCAGCTGATCACCTGGAGAGGACCAAGGGGAGTTTC 1072
 QY 312 LysMetThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysProAsn 331
 DB 1073 AAGATGACGATCCCGACAGAGTGCGCCGCGTGGGAGAGCGGAAGACCAACCCAAC 1132
 QY 332 MetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThr 351
 DB 1133 ATGAACCTACATAGCTACACCGCGCCCTCGTACACTATGACCAAGAACATCATGACC 1192
 QY 352 LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeu 371
 DB 1193 AAGGTCCATGGGAAGCGCTACGCTACAAAGTTGACACTTCCACGGGATCGCCAGGCCCTC 1252
 QY 372 GlnProHisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSer 391
 DB 1253 CAGGCCACCCCGCGAGTATCTCTGTACAAAGTACCCCTACAGCTCCCGTACATGGGC 1312
 QY 392 SerTyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411
 DB 1313 TCTTATACCCGCCACCAAGAGATGAACTTGTGTGGCGCCACCTCCAGGCCCTCC 1372
 QY 412 ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGly 431
 DB 1373 GTGACATCTTCCAGTTTCTTCTGCTCCCAACCCATCTGATGATTCACCACTGGGGGT 1432
 QY 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
 DB 1433 ATATACCCCAACTAGCTCCCGACAGCATATGCTCTCATCTGGGCACTTACTAC 1492

RESULT 4
 : US-09-716-953-1936
 : Sequence 1936, Application US/09716953
 : GENERAL INFORMATION:
 : APPLICANT: Gearling, David P.
 : APPLICANT: Holtzman, Douglas A.
 : APPLICANT: McCarthy, Sean
 : TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 : TITLE OF INVENTION: THEREFOR
 : FILE REFERENCE: 1600.2044-001
 : CURRENT APPLICATION NUMBER: US/09/716.953
 : CURRENT FILING DATE: 2000-11-21
 : PRIOR APPLICATION NUMBER: 60/167,413
 : PRIOR FILING DATE: 1999-11-24
 : NUMBER OF SEQ ID NOS: 2620
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1936
 : LENGTH: 3309
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-716-953-1936

Alignment Scores:
 Pred. No.: 1,31e-161 Length: 3309
 Score: 2197.00 Matches: 410

Percent Similarity: 91.96%
 Best Local Similarity: 89.13%
 Query Match: 28
 Gaps: 4

US-09-902-772-2 (1-451) x US-09-716-953-1936 (1-3309)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 128 ATGGCCAGCAGCTATTAGGAAGCTTATCACTTGTGATGAGGACCGCTGTTTGG 187
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 188 TGTCTCTACGGAAGCCACACCTGCTAGACAGATGACCGGCTCTCTCCAGCAG 247
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 DB 248 TATGGACAGACTTCCAAAGATGAGCCAGCGCTCCCTCAGCAGGATGGCTGTCAAC 307
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 308 CCAGCCAGGGGTACCATCAAAATGATGATGACCTTACCCAGGATGCTTCAAGAAC 367
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 368 TCTCTGATGATGACGATGCTGCGCAAGCGGGGAAAGATGGGGGACAGACCGCTT 427
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThrThr 120
 DB 428 GGGATGACACTAGCGGACCTACATGAGAGAGACATGACACCCCAACATGACACG 487
 QY 121 AsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
 DB 488 AACGAGCGCAGAGTATCGTCCAGCATCTTACGCTATGAGATGACAGCATGTCGG 547
 QY 141 GlnTrpLeuGluTrpAlaValLysGlyLysLeuProAspValAspIleLeuPhe 160
 DB 548 CAGTGGCTGAGTGGGGGAGAAAGATATGCGCTTCAGACGTCACATCTGTATTC 607
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 608 CAGAAATCATGATGGGAAGAACTGTGCAAGATGACCAAGAGCAGCTTCAGAGCTCAC 667
 QY 181 ProSerTyrAsnAlaAspIleLeuSerHisLeuHisTyrLeuArgGluArgGlyAla 200
 DB 668 CCAAGCTTACAGCCGACATCTCTCTCAATCTCCACATGACCAAGAGACTCTCTT 727
 QY 201 ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrArg 220
 DB 728 CCACATTTG-----ACTTCA-----GATGATGTTGATTAAGCTTACAAACTCT 772
 QY 221 Pro-----AspLeuProTyrGluGlnAlaArgSerAla 232
 DB 773 CCACGGTTAAATGCTAGTAAACACAGATTTACCAATATGAGCCGCCAGAGATCAGCC 832
 QY 233 TrpThrSerHisSerHisProThr---GlnSerLysValaThrGlnProSerSerSerThr 251
 DB 833 TGGACCGGTACAGCCGACCCAGCCAGCCAGTCCGTAACCATCTCTCTCCACA 892
 QY 252 ValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThr 271
 DB 893 GTGGCCAAACTGAAAGACAGCGCTCAGTGAATGCTTATCATGATCTTGGACCAACA 952
 QY 272 SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuGlu 291
 DB 953 AGTAGCCGCTTGGAAATCCAGGAGATGGCCAGATCCAGCTTGGCGCTCTCTGAG 1012
 QY 292 LeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyTrpAsnGlyLuphe 311
 DB 1013 CTCTGTGCGACAGCTCCAACTCCAGCTGATCACCCTGGGAGGACCAACGGGAGTTC 1072
 QY 312 LysMetThrAspProAspGluValAlaArgArgTrpGlyLysLeuArgLysSerLysProAsn 331

DB 1073 AAGATGACGAGATCCCGACGAGGTGGCCGCGTGGGGAGAGCGGAAAGCAACCAAC 1132
 QY 332 MetAspTyrAspLysLeuSerArgAlaLeuArgTyrTyrLysAspLysIleMetThr 351
 DB 1133 ATGACATGACATTAAGCTCAGCGCGGCTCGCTGATCTAGTATGACAAAGATCATAC 1192
 QY 352 LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeu 371
 DB 1193 AAGTGCATGGGAAGCGCTACAGCTTACAGTTCATCCAGCGGATGCGCCAGCGCTC 1252
 QY 372 GlnProHisProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSer 391
 DB 1253 CAGCCCAACCCCGCGGAGTATCTCTGTAAGATACCCCTCAGACCTCCGCTCATATGG 1312
 QY 392 SerTyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411
 DB 1313 TCTATACGCGCACCCACAGAGATGATCTGTGGCGCCACCTCCAGCCCTCC 1372
 QY 412 ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLys 431
 DB 1373 GTGACATCTTCCACTTTTTCGTGCGCCCAACCATACATGATGATTCACCACTGGGG 1432
 QY 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
 DB 1433 ATATACCCCAACTAGGCTCCACACGACATGATGCTTCTCATCTGGGCACTTACTAC 1492

RESULT 5

US-09-721-589-5498
 ; Sequence 5498, Application US/09721589
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Villereal, Jean-Luc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: 1600.2045-001
 ; CURRENT APPLICATION NUMBER: US/09/721.589
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR FILING DATE: 1999-11-24
 ; NUMBER OF SEQ ID NOS: 7017
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5498
 ; LENGTH: 3309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-721-589-5498

Alignment Scores:

Align. No.: 1,31e-161 Length: 3309
 Score: 2197.00 Matches: 410
 Percent Similarity: 91.96% Conserved: 13
 Best Local Similarity: 89.13% Mismatches: 23
 Query Match: 29 Indels: 14
 DB: 29 Gaps: 4

US-09-902-772-2 (1-451) x US-09-721-589-5498 (1-3309)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 128 ATGGCCAGCAGCTATTAGGAAGCTTATCACTTGTGATGAGGACCGCTGTTTGG 187
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 188 TGTCTCTACGGAAGCCACACCTGCTAGACAGATGACCGGCTCTCTCTCCAGCAG 247
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 DB 248 TATGGACAGACTTCCAAAGATGAGCCAGCGCTCCCTCAGCAGGATGGCTGTCAAC 307
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 308 CCAGCCAGGGGTACCATCAAAATGATGATGACCTTACCCAGGATGCTTCAAGAAC 367

OY	201	ThrpHelePheProAsnThrSerValLyrProGluAlaThrGlnArgIleThrThrArg	220
Db	728	CCACATTG-----ACTTCA-----GATGATGTTGATTAAGGCGCTTCAAAAACCTCT	772
OY	221	Pro-----AspleuProTyrGluGlnAlaArgArSerAla	232
Db	773	CCACGGTTAAATGCATGCTAGAAACACAGATTATTACATTTGAGCCCCCAGAGATGACCC	832
OY	233	TrpThrSerHisSerHisProThr--GlnSerLysAlaThrGlnProSerSerSerThr	251
Db	833	TGGACCGGTACAGCGCACCGCCACGCGCCACGTGGAAGCTGCTCAACCATCTCTCCACAA	892
OY	252	ValProLysThrGlnLysPglNArgProGlnLeuAspProTyrGlnIleLeuGlyProThr	271
Db	893	GTGGCCCAAACTGGAAGACACCGCTCTCAGTTAGTCTTACGATTTCTTGAGCAACA	952
OY	272	SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPglNPhenLeuGln	291
Db	953	AGTACGCCCTTGGCAAAATCCAGGCAGTGGCCAGATCCAGCTTTGGCACTTCTCTCGAG	1012
OY	292	LeuLeuSerAspSerSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLeuPhe	311
Db	1013	CTCCTGTGTGGACAGCTCCAACTCCAGCTGCATCACTGTGGAAAGGACCAAGGGGAGTTC	1072
OY	312	LysMetThrAspProAspGlnValAlaArgArgTrpGlyGluThyGlySerLysProAsn	331
Db	1073	AAGATGACGGATCCCGACGAGGTGGCCGCGCGTGGGAGAGCGGAAAGCAAAACCCAAAC	1132
OY	332	MetAsnTyrAspLysLeuSerSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThr	351
Db	1133	ATGAACTACGATTAAGCTCAGCCGCGCCCTCCGTTACTACTATGACAAAGAAATCATCTGACC	1192
OY	352	LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeu	371
Db	1193	AAGGTCACTGGGAAGCGGTACGCTCAAGATTGCATTCACAGGGATGGCCAGGCCCTC	1252
OY	372	GlnProHisProProGlnUserSerMetLysTyrLysTyrProSerAspLeuProTyrMetSer	391
Db	1253	CAGGCCCAACCCCCGGAGTCACTCTGTACAAAGTAAACCCCTCAGACCTCCCTACATGGGC	1312
OY	392	SerTyrHisAlaHisProGlnLysMetAsnPhenValAlaProHisProProAlaLeuPro	411
Db	1313	TCTCATCAGCGCCACCCACAGAAATGAACTTTGGGGGCCCAACCTCCAGCCCTCCCC	1372
OY	412	ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGly	431
Db	1373	GTGACACTCTCCAGATTTTTTTGGCGCCCAAAACCATCTGGAATTCCACCACTGGGGGT	1432
OY	432	IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	451
Db	1433	ATAATCCCAACACTAGAGCTCCCAACAGCCATATGCTCTCATCTGGGACATTACTAC	1492
RESULT 7			
US-09-726-811-4920			
Sequence 4920, Application US/09726811			
GENERAL INFORMATION:			
APPLICANT: Gutierrez-Ramos, Jose-Carlos			
APPLICANT: Welch, Nadine S.			
APPLICANT: Wen, Danyi			
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES			
TITLE OF INVENTION: THEREFOR			
FILE REFERENCE: 1600.2027-001			
CURRENT FILING DATE: 2000-11-30			
PRIOR APPLICATION NUMBER: 60/168,136			
PRIOR FILING DATE: 1999-11-30			
NUMBER OF SEQ ID NOS: 5515			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 4920			
LENGTH: 3309			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-726-811-4920			

Alignment Scores:			
Score:	1,31e-161	Length:	3309
Percent Similarity:	91.96%	Matches:	410
Best Local Similarity:	89.13%	Mismatches:	13
Query Match:	89.89%	Indels:	14
DB:	29	Gaps:	4
US-09-902-772-2 (1-451) x US-09-726-811-4920 (1-3309)			
QY	1	MetaIaSerThrIleLysGluAlaLeuSerValIaSerGluAspGlnSerLeuPheGlu	20
DB	128	ATGGCCACACATATTAAAGAACCCCTTACATGTTGGAGGAGACACAGCTGTTTGAG	187
QY	21	CysAlaIaYrGlySerProHisLeuAlaYsthrGluMetThrAlaSerSerSerGlu	40
DB	188	TGTGGCTTCAGGAACCCACACCTGGCTTAAGACAGATGACCGCTCCCTCCAGCGAC	247
QY	41	TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro	60
DB	248	TATGGACAGACTCCAAAGATGAGCCACGCGCTCCCTCAGACAGATTGGCTGTCAACCC	307
QY	61	ProAlaArGValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn	80
DB	308	CCAGCAGAGGTGCACCATCAAAATGGAATGATACCTAGCCAGAGTGAATGCTCAAGAAC	367
QY	81	SerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnVal	100
DB	368	TCTCTGTATGATGATGACGTGTGGCCAAAGCCGGAAGATGCTGGGACGCCACACACCGTT	427
QY	101	GlyMetAsnYrGlySerTyrMetGluGlnYsthrIleProProProAsnMetThrThr	120
DB	428	GGGATGAACCTACGGACGCTACTGTGAGGAGAGACCATGCCACCCCAACATGACACAG	487
QY	121	AsnGluArGArGValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg	140
DB	488	AACGAGCCAGAGTTATGTCGACACAGATCCTACGATATGAGATACAGACCATGTGGCG	547
QY	141	GlnTyrPLeuGlnTyrPLeuValIysGlnTyrGlyLeuProAspValAspIleLeuLeuPhe	160
DB	548	CAGTGGCTGGAGTGGCGGTGAAMAATATGGCTCTCCAGACGTCAACCTCTGTTATTC	607
QY	161	GlnAsnIleAspGlyLysGlnLeuGlysthrLysMetThrLysAspAspPheGlnArgLeuThr	180
DB	608	CAGAACATCGATGGAGAGAACTGTGCAGATGACCAAGACGACTTCCAGAGGCTCACCC	667
QY	181	ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuAlaArgIleuArgIleuAla	200
DB	668	CCGAGCTCAACAGCCGACATCCTCTCCACATCTCCACTCCTCAGAGACATCCTCTCT	727
QY	201	ThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArg	220
DB	728	CCACATTTTG-----ACTTCA-----GATGAGTGTGATAAAGCCCTTACAAAACTCT	772
QY	221	Pro-----AspLeuProTyrGluGlnAlaArgArgSerAla	232
DB	773	CCACGGTTAAATGCATGCTAGAAACACATTTACATATATGACCCCCCAGAGATACAGCC	832
QY	233	TyrThrSerHisSerHisProThr--GlnSerLysAlaThrGlnProSerSerSerThr	251
DB	833	TGGACCGGTACAGGCCACCCACGCCCGCCAGTGAAGAGCTGCACACCATCTCTCCACCA	892
QY	252	ValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThr	271
DB	893	GTGGCCAAATCTGAAGACACAGCTCTCAGTTAGATCCTTATCAATATCTTGGACCAACA	952
QY	272	SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlu	291
DB	953	AGTAGCCGCCCTTGCATAATCCAGGACAGTGGCCAGATCCAGCTTTGGCACTTCTCTGGAG	1012
QY	292	LeuLeuSerAspSerSerAsnSerAsnGlyIleThrTyrGluGlyThrAsnGlyLeuPhe	311

Db 1013 CTCCTGTCGAGACGCTCCAACTCCAGCTGCATTCACCTTGGAGGACCAAGGGAGATTTC 1072
Qy 312 LysMetThrAspProAspGluValAlaIArgArgTrpGlyGluArgLysSerLysProAsn 331
Db 1073 AAGATGACGATGCCGACGAGAGTGGCCGCGCTGGGAGAGCGGAAGCAAAACCCAAAC 1132
Qy 332 MetAsnTrpAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnLleMetThr 351
Db 1133 ATGAACCTACGATGAGCTAGCGCGCCCTCCGTTACTATGACAAAGATCATGTAGCC 1192
Qy 352 LysValHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyLleAlaGlnAlaLeu 371
Db 1193 AAGGCTCATGGAGAGCGCTAGCCTACAGCTTCACACTTCCAGGGATGCCAGGCCCTC 1252
Qy 372 GlnProHisProProGlnLysSerSerMetLysTrpTrpProSerAspLeuProTyrMetSer 391
Db 1253 CAGCCGCCACCCCGGAGATCATCTCTGTACAGTACCCTCAGACTCCCTCATCTGCGC 1312
Qy 392 SerTrpHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuPro 411
Db 1313 TCCTATACGCCACCCACAGAGATGACTTTGGGGCCGCCACCTCCAGCCCTCC 1372
Qy 412 ValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGly 431
Db 1373 GTGACATCTTCAGATTTTTGGCTGCCCAACCCATCTGGAATTCACCAACTGGGGGT 1432
Qy 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTrp 451
Db 1433 ATATACCCCAACACTAGGCTCCCGCAGCATATGCTCTCATCTGGGCACTTACTAC 1492

RESULT 8
US-09-948-941-233
; Sequence 233, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO00788
; CURRENT APPLICATION NUMBER: US/09/948, 941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231, 328
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 233
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-233

Alignment Scores:
Pred. No.: 6,89e-161 Length: 1389
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
Gaps: 4

US-09-902-772-2 (1-451) x US-09-948-941-233 (1-1389)

Qy 2 AlAserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLysPheGluLys 21
Db 25 GCACCTCATATCAAGGAAGCCCTTATGAGTGTGAGTGGACCAAGTCGTTTGTGAGTGT 84
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyLys 41
Db 85 GCCACGGAAGCAACACCTGCTAAGACAGATGACCGGCTCTCTCTCCAGGACTAT 144
Qy 42 GlycineThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db 145 GACAGACTTCCAGAGTAGAGCCACGCGTCCCTCAGCAGATGGCTGTCTCAACCCCA 204
Qy 62 AlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsnSer 81

Db 205 GCCAGGGTCACCATCAAAATGAAATGTAACCTTACCCAGGAGGAAATGCTCAAGAAACTCT 264
Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 265 CCTGATGATGACAGTGTGCGCAAAAGCGGGAAGATGTGGGGACACCCAGAACCGCTTGG 324
Qy 102 MetAsnTrpGlySerTrpMetGluLysHisIleProProProAsnMetThrAsn 121
Db 325 ATGAACCTACGCGACCTACATGAGAGACCAATGTCACCCCAAAACATGACCACCAAC 384
Qy 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerTrpAspHisValArgGln 141
Db 385 GAGCGAGAGTTATGTCGACAGCATCTCAGTATGAGATGAGACCAATGTCGCGGAG 444
Qy 142 TrpLeuGluTrpAlaValLysGlyLysTrpLysLeuProAspValAspIleLeuLeuPheGln 161
Db 445 TGGCTGAGAGTGGGGGTGAAGAAATATGCTTCACAGACGTCAACATCTTGTATTCAG 504
Qy 162 AsnIleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCGATGGGAAGAACTGTGCAAGATGACCAAGAGCACTTCCAGAGCTCACCCCC 564
Qy 182 SerTrpAsnAlaAspIleLeuSerHisIleuHisTrpLeuArgGluArgGlyAlaThr 201
Db 565 AGCTACAAAGCCGACATCTTCTCATATCTCAGTACCTCAGAGAGACTCTCTTCCA 624
Qy 202 PheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrTrpArgPro 221
Db 625 CATTTG-----ACTTCA-----GATATGTTGATAAAGCTTACAAAACTCTCCA 669
Qy 222 -----AspLeuProTyrGluGlnAlaIArgSerAlaTrp 233
Db 670 CGGTTAATGATGCTAGAAACACACGATTTTACATATGAGCCCCAGAGAGATCAGCCTGG 729
Qy 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db 730 ACCGGTACAGGCGCACCCACGCGCCAGTGCAGAACTGCTCAACATCTCTCCACAGTG 789
Qy 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db 790 CCCAAACTGAAAGACGACGCTCAGATTAATGCTTATGATGATTTGGACCAACAGT 849
Qy 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeu 292
Db 850 AGCCGCTTGCAAAATCCAGGACGAGTGCAGATCCAGCTTGGCAGTTCCCTCGAGCTC 909
Qy 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLys 312
Db 910 CTGTGGACACTCACTCACTGCTGATCAGCTGGGAGAGACCAACGGGAGTTCAAG 969
Qy 313 MetThrAspProAspGluValAlaIArgArgTrpGlyGluArgLysSerLysProAsnMet 332
Db 970 ATGACGATGCCGACGAGAGTGGCCGCGCGTGGGAGAGCGGAAGCAAAACCAACATG 1029
Qy 333 AsnTrpAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnLleMetThrLys 352
Db 1030 AACTACGATTAAGCTCACCGCGCCCTCGTTACTATGATGACAAAGCAATCATATACCAAG 1089
Qy 353 ValHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyLleAlaGlnAlaLeuGln 372
Db 1090 GTCCATGGGAAGCCCTACGCTTCAAGATGTCAGACTTCCACGGGATGCCAGGCCCTCAG 1149
Qy 373 ProHisProProGlnLysSerMetLysTrpTrpProSerAspLeuProTyrMetSerSer 392
Db 1150 CCCACACCCCGGAGATCATCTCTGTAAGATACCCCTCAGACCTCCGCTATAGGGCTCC 1209
Qy 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db 1210 TATACGCCCAACCAAGAGATGAACTTGTGGCGCCCAACCTCCAGGCTCCCGGTG 1269
Qy 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle 432


```

Db      1270 ACATCTTCCAGTTTCTTCTGCCCAAAACCCATGCAATTCACCACTGGGGGTATA 1329
OY      433 TTPProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      1330 TACCCCAACACTAGGCTCCCAACAGCCATATGCTTTCATCTGTGGCACTACTAC 1386

RESULT 9
PCT-US02-23913-98
; Sequence 98, Application PC/TUS0223913
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-04APC
; CURRENT APPLICATION NUMBER: PCT/US02/23913
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-23913-98

Alignment Scores:
Pred. No.:      1,86e-160      Length:      3166
Score:          2182.00      Matches:      407
Percent Similarity: 91.72%      Conservative: 14
Best Local Similarity: 88.67%      Mismatches: 24
Query Match:    89.28%      Indels:      14
Db:             1             Gaps:         4

US-09-902-772-2 (1-451) x PCT-US02-23913-98 (1-3166)
OY      2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlyCys 21
Db      281 GCAGCTCATATCAAGAGAGAGGCTTATCATGTTGAGAGAGGACAGTCGTTGAGAGT 340
OY      22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlyTyr 41
Db      341 GCTTACGGAAGCGCACACCTGCTTAAGACAGAGATGACCGCTCTCTCCAGCGACTAT 400
OY      42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db      401 GGACAGACTTCCAAATGAGCCACCGCGTCCACAGAGATTGGCTGTCTCAACCCCA 460
OY      62 AlaArgValThrIleLysMetGlyCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db      461 GCCAGGGTCAACATCAAAATGGAATTAACCCCTAGCAGGTGAATGGCTCAAGAACTCT 520
OY      82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGly 101
Db      521 CCGATGATGATGCGAGTGTGGCCAAAGCGGGAGAGATGTGTGGCCAGCCAGACCGTTGG 580
OY      102 MetAsnTyrGlySerTyrMetGluGlyLysHisIleProProProAsnMetThrThrAsn 121
Db      581 ATGAAGTACGCGAGCTACATGAGGAGAGACATCCACCCCAAAACATGACACGAGAAC 640
OY      122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db      641 GACCGCAGAGTATCTGCGCAGCAGAGATCTACGCTATGAGATACAGACCATGTGCGCAG 700

OY      142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db      701 TGGCTGGAGTGGCGCGTGAAGAAATATGCTTCCAGACTCATCTTGTATTCCAG 760
OY      162 AsnIleAspLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      761 AACATGATGGAGGAAGAACTGTGCAGATGACCAAGAGACACTTCCAGAGACTCACCCC 820
OY      182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db      821 AGCTAACAGCCGACATCTCTCTCAATCTCCATCTCCAGAGAGACTCTCTTCCA 880
OY      202 PheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgGlyLeuThrArgPro 221
Db      881 CATTTG-----ACTTCA-----GATGATGTGATTAACCTTACAAACTCTCCA 925
OY      222 -----AspLeuProTyrGlnGlnAlaArgArgSerAlaTrp 233
Db      926 CGGTTAATGATGCTGTAAGAAACACAGATTTACCATATGAGCCCCCAGAGATCAGCCTGG 985
OY      234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db      986 ACCGGTCAGCGCCACCCACAGCCCGCCAGTGAAGAGCTTCAACCATCTCTTCCACAGTG 1045
OY      253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db      1046 CCCAAGACTGAACACAGCGCTCTCAATTAATGATCTTATGATTCCTTGAGCCACACAAAGT 1105
OY      273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 292
Db      1106 AGCCGCTTGCAAAATCCAGGCACTGGCCAGATCCAGATCCACTTTGGCAGTTCCTCGAGCTC 1165
OY      293 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLeuPheLys 312
Db      1166 CTGTGCGACAGCTCCAACTCCAGCTCATCTCATCTGGGAAGGACCAACCGGGAGTTCAAG 1225
OY      313 MetThrAspProAspGlnValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
Db      1226 ATACGCGATGCCAGACAGTGGCCCGCGCTGGGGAGAGCGGAGAGCAACCAACATG 1285
OY      333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db      1286 AACTACAGATPAGCTCAGCGCGCGCTCCGTTACTACTATGACAAAGAACATCATGACCAAG 1345
OY      353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
Db      1346 GTCCATGGGAAGGGCTACGCTTAAAGTTGCACTTCCAGGGGATCCCGACGGCCCTCCAG 1405
OY      373 ProHisProProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392
Db      1406 CCCCACCCCGCGAGTCACTCTGTACAAAGTACCCCTTCAGACTCCCGTACATGGGCTCC 1465
OY      393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db      1466 TATCAGCGCCACCAAGAAATGAACTTTGTGGCCGCCACCCCTCCAGCGCCCGCG 1525
OY      413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle 432
Db      1526 ACATCTTCCAGTTTCTTCTGCGCCCAAAACCTACTGGAATTCACCAACTGGGGGTATA 1585
OY      433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db      1586 TACCCCAACACTAGGCTCCCAACAGCCATATGCTTTCATCTGTGGCACTACTAC 1642

RESULT 10
US-09-784-356-26
; Sequence 26, Application US/09784356
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,

```

```

; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; FILE REFERENCE: Modulators
; CURRENT APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-356-26

Alignment Scores:
Pred. No.: 1,86e-160 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 30 Gaps: 4

US-09-902-772-2 (1-451) x US-09-784-356-26 (1-3166)
QY 2 AlasethrilleysglualaaleuSerValSerGlunspInserleupheglucys 21
DB 281 GCAGCTCATATCAAGAGACCTTATCATGTTGAGTGAGAGACCAAGCTGTTGAGTGT 340
QY 22 AlatyrglyserProhIsleuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41
DB 341 GCTACGAGAACGCCACCTGCTAAGACAGATGACCCGTCCTCCCTCAGAGACTAT 400
QY 42 GlyGluThrserlysmetSerProargValProGlnGlnAspThrleuSerGlnProPro 61
DB 401 GGACAGACTTCCAGAGAGACCCACGCGTCCCTCAGCAGAGATTGGCTGTCACCCCA 460
QY 62 AlarGValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
DB 461 GCAGAGGTCAACCAAAAGATGTAACCTAGCCAGGTGATGCTCAAGAACTCT 520
QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
DB 521 CCGATGAGAAATGAGTGTGGCCAAAGCGGAGAAAGATGTTGGGACCCAGACCGTTGGG 580
QY 102 MetAsnTyrglyserIlyMetGluLysHisIleProProProAsnMetThrThrAsn 121
DB 581 ATGACTACGCGCAGCTACATGAGAGAAACACATGCAACCCCAACATGACACGAAAC 640
QY 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
DB 641 GACGCGAGATTTCTGTCAGCAGATCCTACGCTATGAGTACAGACCACTGTGCGGAG 700
QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
DB 701 TGGCTGAGAGTGGCGGTGAAGAAATATGGCTTCCAGACGCTCAACTCTGTATTCCAG 760
QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
DB 761 AACATGAGTGGAGAGAACTGTCAAGATGACCAAGAGACACTTCCAGAGGCTCACCCCC 820
QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArgGlyAlaThr 201
DB 821 AGCTACAGCCCAACATCTTCTCTCAGATCTCCATCTCAGAGAGACTCCTCTTCCA 880
QY 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro 221
DB 881 CATTTG-----ACTTCA-----GATGATGTTGATTAAGCCTTACAAACTCTCCA 925
QY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTrp 233

```

```

DB 926 CGGTTAATGATCTTGAAACACAGATTACCATATGAGACCCCCAGAGATAGCCTGG 985
QY 234 ThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerThrVal 252
DB 986 ACCGGTCAGCGGCACCCACAGCCAGGTGAGAAAGCTGCTCAACATCTCTTCCACAGTG 1045
QY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
DB 1046 CCCAAACTGMAACACAGCGCTCTCAGTTAGATCTTATCAGATCTTGGACCAACAGT 1105
QY 273 SerArgLeuAlaAsnProGlnSerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 292
DB 1106 AGCCGCTTGCAATATCCAGGACAGTGGCCAGATCCAGATCTTGGCAGTTCTCTGGAGCTC 1165
QY 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLys 312
DB 1166 CTGTCGAGACAGCTCCAACTCCAGCTCAGTCTGAGTGGGAGGCCAACAGGGAGTTCAAG 1225
QY 313 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
DB 1226 ATGACGAGATCCGACAGAGTGGCGCGCTGGGAGAGCGGAAAGACCAACCAACATG 1285
QY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
DB 1286 AACTACGATTAAGCTCAGCGCGCGCTCCCTTACTACTATGACAAAGACATCATGACCAAG 1345
QY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
DB 1346 GTCCATGAGGAGAGGCGTACGCTCAAGTTGATTCACGAGATGCGCCAGCGCTCCAG 1405
QY 373 ProHisProProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392
DB 1406 CCCACCCCGGAGATCATCTGTCAAGATACCTCCAGACCTCCGATACATAGGCGCTGC 1465
QY 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
DB 1466 TATCAGCCCGCACCAAGAGATGAACTTGTGCGCCCGACCTCCAGCGCTCCCGGAG 1525
QY 413 ThrSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyLe 432
DB 1526 ACATCTTCCAGTTTGTGTGTCGCCAAACCCATCTGGAATTCACCAACTGGGGGTATA 1585
QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
DB 1586 TACCCCAACACTGAGCTCCCGCACAGCATATGCTTCTCATCTGCGGCACTTACTAC 1642

RESULT 11
US-10-021-660-26
; Sequence 26, Application US/10021660
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-26

Alignment Scores:

```


Pred. No.: 1,86e-160 Length: 3166
 Score: 2182.00 Matches: 407
 Percent Similarity: 91.72% Conservative: 14
 Best Local Similarity: 88.67% Mismatches: 24
 Query Match: 89.28% Indels: 14
 Gaps: 4

US-09-902-772-2 (1-451) x US-10-021-660-26 (1-3166)

QY 2 AlaSerThrIleIysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
 Db 281 GCACCTATATCAAGGAGGAGCCTTATCACTGTGTGAGTGAGGACAGTCTTTTGAGAGT 340
 QY 22 AlaTyrGlySerProHisIleuAlaIysThrGluMetThrAlaSerSerSerGlyTyr 41
 Db 341 GCCTACGAGACGCCACACCTGGCTAAGACAGATGACGCCGCTCTCTCCAGGACAT 400
 QY 42 GlyIleThrSerIysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
 Db 401 GGACAGACTTCCAGATGAGCCACGCGTCCCTCAGCAGATGGCTGTCTCAACCCCA 460
 QY 62 AlaArgValThrIleIysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 Db 461 GCCAGGGTCACCATCAAAATGGAATGTAACTTGAACCTGAGTGAATGGCTCAAGAACTCT 520
 QY 82 ProAspAspCysSerValAlaIysGlyIysMetValSerSerSerAspAsnValGly 101
 Db 521 CCTATGATCATCTAGTGGCCAAAGCGGAGAGATGGTGGGACGCCAGACACCGTTGGG 580
 QY 102 MetAsnTyrGlySerTyrMetGluGluIysHisIleProProProAsnMetThrAsn 121
 Db 581 ATGAACTACAGCGACCTCATGTGAGAGAGACATGCGACCCCAACATGACCAAGAAC 640
 QY 122 GluArgTrpValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
 Db 641 GAGGCGAGAGTTATGTCGACGAGATCTTACGCTATGAGATTAAGACACAGTGGCGAG 700
 QY 142 TrpLeuGluTrpAlaValIysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
 Db 701 TGGGTGAGTGGCGGCTGAAGATATGCTTCACAGACGTCACATCTTTTATCTCAG 760
 QY 162 AsnIleAspGlyIysGluLeuCysIysMetThrIysAspAspPheGlnArgLeuThrPro 181
 Db 761 AACATCGATGGAGGAGAGCTGTGCAAGATGACCAAGAGCACTTCCAGAGCTCACCC 820
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArgGlyAlaThr 201
 Db 821 AGCTACACAGCGGACATCTTCTCTCATCTCCACATCACTCAAGAGACCTCTTCCA 880
 QY 202 PheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArgPro 221
 Db 881 CATTTG-----ACTTCA-----GATGATGTGTATTAAGCCCTTCAAAACTCTTCA 925
 QY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTrp 233
 Db 926 CGGTAAATGATGCTAGAAACACAGATTACCATATAGAGCCCAAGAGATGAGCCCTGG 985
 QY 234 ThrSerHisSerHisProThr---GlnSerIysAlaThrGlnProSerSerThrVal 252
 Db 986 ACCGATACAGGACCCACCCAGCCAGTCGAAAGCTGCTCAACCATCTCTTCCACAGT 1045
 QY 253 ProIysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
 Db 1046 CCCAAACTGAAGACAGCGTCTCTCATGATTGATCTTATGATCTTTGAGCCAAACAGT 1105
 QY 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeu 292
 Db 1106 AGCGGCTTGCAGAAATCCAGAGAGTGGCCAGATCCAGCTTGGAGTTCTCTGGAGCTC 1165
 QY 293 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheIys 312
 Db 1166 CTGTGCGACAGCTCCAACTCAGCTGATACACTGCGGAGGAGGACCAAGGGAGTTCAMG 1225

QY 313 MetThrAspProAspGluValAlaArgArgTrpGlyIleArgIysSerIysProAsnMet 332
 Db 1226 ATGAGGATTCGCCAGAGAGTGGCGGCTGGGAGAGCGGAAGAGCAAAACCAACATG 1285
 QY 333 AsnTyrAspIysLeuSerArgAlaLeuArgTyrTyrTyrAspIysAsnIleMetThrIys 352
 Db 1286 AACTACGATTAAGCTCAGCGCGGCGCTCCGTACTACTATGATGACAAAGATCATATGACCAAG 1345
 QY 353 ValHisGlyIysArgTyrAlaTyrIlyIysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
 Db 1346 GTCCATGGGAGAGCGCTACGCTACAGATTCACAGGAGATGCCAGGCGCTCCAG 1405
 QY 373 ProHisProProGluSerSerMetTyrIysTyrProSerAspLeuProTyrMetSerSer 392
 Db 1406 CCCAGCCCCGGAGTATCTCTGTACAGATACCCCTCAGACCTCCGTATATGGGCTCC 1465
 QY 393 TyrHisAlaHisProGlnIysMetAsnAspPheValAlaProHisProProAlaLeuProVal 412
 Db 1466 TATCAGCGCCACCCACAGAAAGATGAACCTTGTGGCGGCCACCTCCAGCCCTCCGCTG 1525
 QY 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIle 432
 Db 1526 ACATCTTCCAGTTTTTTTGTGTCGCCCAAAACCATACCTGAATTCACCAACTGGGGATTA 1585
 QY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
 Db 1586 TACCCCAACACTAGGCTCCACACGACCATATGCTTCTCATCTGGGACTTACTAC 1642

RESULT 12
 US-10-205-823-98
 ; Sequence 98, Application US/10205823
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Gannavarapu, Manjula
 ; APPLICANT: Gorbacheva, Bella
 ; APPLICANT: Hoersch, Sebastian
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Womsey, Angela M.
 ; APPLICANT: Glatt, Karen
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Anderson, Dustin
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; FILE REFERENCE: MRI-044
 ; CURRENT APPLICATION NUMBER: US/10/205,823
 ; CURRENT FILING DATE: 2002-07-25
 ; PRIOR APPLICATION NUMBER: 60/307,982
 ; PRIOR FILING DATE: 2001-07-25
 ; PRIOR APPLICATION NUMBER: 60/314,356
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/325,020
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: 60/341,746
 ; PRIOR FILING DATE: 2001-12-12
 ; PRIOR APPLICATION NUMBER: 60/362,158
 ; PRIOR FILING DATE: 2002-03-05
 ; NUMBER OF SEQ ID NOS: 455
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 98
 ; LENGTH: 3166
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-205-823-98

Alignment Scores:
 Pred. No.: 1,86e-160 Length: 3166
 Score: 2182.00 Matches: 407
 Percent Similarity: 91.72% Conservative: 14
 Best Local Similarity: 88.67% Mismatches: 24
 Query Match: 89.28% Indels: 14


```

Oy 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
    |||||
Db 467 AATGAGGCGAGAGATGATCGTCCCGAGATGCTCTGTGGAGACAGACAGCATGTCGA 526
Oy 141 GlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPhe 160
    |||||
Db 527 CAGGGCTGGGTGGGGGTGAAGAAATATGGCTCTCGATGGAGCGTCTTACTATT 586
Oy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
    |||||
Db 587 CAGATATCGATGGAGAGAGCTGTGCAGATGACAAAGAGTGAAGTCTTCCAGCGGCTCAG 646
Oy 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg----- 198
    |||||
Db 647 CCGAGCTACATGCCGAGATTTCTCTCATCTCCACATCCAGACAGAGACGCCCTT 706
Oy 198 ----- 198
Db 707 CCACATCTGACTTCGATGACGTTGATTAAGGCTTTACAAAACCTCCACGGTTATGCA 766
Oy 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213
    |||||
Db 767 GCCAGAAACACAGGGGGTGAAGCTTTATTTCCCAATCTTCAATATCCCGAAGCT 826
Oy 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 233
    |||||
Db 827 ACGCAAGAAATTAACACTAG----- 847
Oy 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 253
    |||||
Db 847 ----- 847
Oy 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
    |||||
Db 847 ----- 847
Oy 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGluLeuLeu 293
    |||||
Db 848 -----CCAGGTAGGGCGCAATCCAGCTGTGGCAGTTCCGTCCGAACTCTG 895
Oy 294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMet 313
    |||||
Db 896 TCANACAGCTCCAACTCAATGCTCATCTGCGGAGGAGCAACCAAGGGAGTTCAAGATG 955
Oy 314 ThrAspProAspGluValAlaArgArgTrpGlyLysArgLysSerLysProAsnMetAsn 333
    |||||
Db 956 ACAGACCCGAGACGAGGTGGCTGGCGGTGGGGAGAGAGCAAGCCCAACATGTAAC 1015
Oy 334 TyrAspLysLeuSerAlaAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
    |||||
Db 1016 TATGACAAGCTCAGCGCGCGCTCCGCTACTACTACACAAACATCATGACCAAGGTG 1075
Oy 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
    |||||
Db 1076 CACGGGAGAGCGCTACGCTCAAGTTGACTTCCACGGAGATGGCCAGCGCTGCACAGCC 1135
Oy 374 HisProProGlnUserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
    |||||
Db 1136 CACCTCTGATGCTGCTCTGATACAGTACCCCTCCAGCTCCATACATGAGGCTCTAT 1195
Oy 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
    |||||
Db 1196 CAGGCCACCCCGAGAGATGAATTTGTGTCCCGACCTCCGCTCCGCTCCAGTCCACA 1255
Oy 414 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyr 433
    |||||
Db 1256 TCTTCCAGTTCTTCTTCCCGCAACCCATACGGAATTCACGAGATGGGGGAGATCTAC 1315
Oy 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
    |||||
Db 1316 CCGAACACTAGGCTCCAGCCAGCCATATGCTCTCAGCTGGGCACTTACTAC 1369

```

```

US-09-442-384A-475
; Sequence 475, Application US/09442384A
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Lukashiev, Malvey
; TITLE OF INVENTION: Hematology/Immunology Array
; FILE REFERENCE: CLON-006CIP15
; CURRENT APPLICATION NUMBER: US/09/442,384A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 09/053,375
; PRIOR FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 830
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 3126
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-442-384A-475

Alignment Scores:
Pred. No.: 2,72e-131 Length: 3126
Score: 1810.00 Matches: 336
Percent Similarity: 90.34% Conservative: 10
Best Local Similarity: 87.73% Mismatches: 23
Query Match: 74.06% Indels: 14
DB: 18 Gaps: 4

US-09-902-772-2 (1-451) x US-09-442-384A-475 (1-3126)
Oy 78 SerArgAsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSer 97
    |||||
Db 150 TCAAGGAATCTCTCTGATGAATGACGTGTGCCAAAGCGGAAAGATGTGGGACGCCA 209
Oy 98 AspAsnValAlaLysMetAsnTyrGlySerTyrMetGluGlyLysHisIleProProAsn 117
    |||||
Db 210 GACACCGTTGGGTGAATCACTACGCGAGCTACATGAGAGAGACACATGCCACCCCAAC 269
Oy 118 MetThrAsnGlnArgValIleValProAlaAspProThrLeuTrpSerThrAsp 137
    |||||
Db 270 ATGACCCGAAACGAGCCGCAAGATTATGTGCCAGACATCTTACGCTATGAGTACAGAC 329
Oy 138 HisValArgIleTrpLeuGluTrpAlaValLysGluLysGlyLeuProAspValAspIle 157
    |||||
Db 330 CATGTGGGAGAGTGGCTGGAGTGGCGGTGAAGAAATATGCGCTCCAGACGTCACATC 389
Oy 158 LeuLeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGln 177
    |||||
Db 390 TTGTTATTCAGAAACATCGATGGAGAGAACTGTGCAGATGACCAAGAGCACTTCCAG 449
Oy 178 ArgLeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlu 197
    |||||
Db 450 AGGCTCACCCCGAGCTCAAGCCGAGATCTTCTCTCAGATCTCCAGTCTCAGAGAG 509
Oy 198 ArgGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIle 217
    |||||
Db 510 ACTGCTCTCCACATTTG-----ACTTCA-----GATGATGTGTAATAAGCCTTA 554
Oy 218 ThrThrArgPro-----AspLeuProTyrGluGlnAlaArg 229
    |||||
Db 555 CAAAACTCTCCAGGTTAATGATGCTAGAAACACAGATTTACCATATGAGGCCCCAGG 614
Oy 230 ArgSerAlaThrPheSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSer 248
    |||||
Db 615 AGATCAAGCTTGAGACCGGTTCACGGCCACCCACGCCCGCAGTGCAGAAAGCTGTCAACATCT 674
Oy 249 SerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeu 268
    |||||
Db 675 CTTTCCAGATGGCCAAAACTGAAGACACAGCGTCTCAGTTAATGCTTATCAGATTTCT 734
Oy 269 GlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPhe 288
    |||||
Db 735 GGACCAACAGATGAGCGCTTGGCAAAATCCAGGAGTGGCCAGATCCAGCTTGGCAGTTTC 794

```

RESULT 14

Wed Jul 30 13:29:52 2003

us-09-902-772-2.rnp

Page 15

Search completed: July 28, 2003, 09:24:09
Job time : 3002.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:30:39 ; Search time 1030.65 Seconds
(without alignments)
3233.216 Million cell updates/sec

Title: US-09-902-772-2
Perfect score: 2444
Sequence: 1 MASTIKALSVEEDSLFE.....IYPNTRLPAAHMPSHLCTYY 451

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Xgapop 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searched: 8407786 segs, 3694357880 residues
Total number of hits satisfying chosen parameters: 16815572

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-x1h
-O/cgnt2_1/USPTO.spool/US09902772/rnatc_23072003_093658_14943/app_query.fasta.1.1294
-DB-Pending-Patents_NA.New -QEXT-fastcap -SUFFIX-trpna -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cd1
-LIST=45 -DOCUMENT=200 -THR.SCORE-pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE-LOCAL -OUTEXT-plo -NOM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09902772.ecgn.1.1.325.etrnat.23072003.093658.14943 -NCPU=6 -ICPU=3
-NO_MMAP -IAREGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgnt2_6/ptodata/2/pna/PCM_NEW_COMB.seq:*
2: /cgnt2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
3: /cgnt2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
4: /cgnt2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
5: /cgnt2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
6: /cgnt2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
7: /cgnt2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
8: /cgnt2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
9: /cgnt2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
10: /cgnt2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
11: /cgnt2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
12: /cgnt2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
13: /cgnt2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
14: /cgnt2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
15: /cgnt2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
16: /cgnt2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
17: /cgnt2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
18: /cgnt2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
19: /cgnt2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
20: /cgnt2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	2298	94.0	1930	15	US-10-170-235-28132	Sequence 28132, A
2	2298	94.0	1930	15	US-60-452-680-11396	Sequence 11396, A
3	2197	89.9	3294	18	US-10-170-235-28519	Sequence 28519, A
4	2197	89.9	3294	18	US-60-452-680-11397	Sequence 11397, A
5	2182	89.3	1389	2	PCT-US02-04915-41	Sequence 41, Appl
6	2182	89.3	1389	14	US-10-211-662-41	Sequence 41, Appl
7	2182	89.3	1389	15	US-10-087-192-1127	Sequence 1127, Ap
8	2182	89.3	1389	15	US-10-126-052A-330	Sequence 330, App
9	2057	84.2	1297	15	US-10-087-192-1124	Sequence 1124, App
10	1991	81.5	2172	14	US-10-144-771-355	Sequence 355, App
11	1624.5	66.5	1359	2	PCT-US02-41414-1210	Sequence 1210, App
12	1624.5	66.5	2957	2	PCT-US02-41414-1209	Sequence 1209, App
13	1624.5	66.5	2957	14	US-10-342-887-710	Sequence 710, App
14	1613.5	66.0	1359	2	PCT-US02-41414-1207	Sequence 1207, App
15	1613.5	66.0	1729	14	US-10-144-771-10845	Sequence 10845, A
16	1613.5	66.0	1729	14	US-10-170-235-34391	Sequence 34391, A
17	1610.5	65.9	2769	15	US-09-949-016-4688	Sequence 4688, Ap
18	1607.5	65.8	2172	11	US-10-087-192-1126	Sequence 1126, Ap
19	956.5	39.1	212231	15	US-09-947-914-53	Sequence 53, Appl
20	956.5	39.1	8059021	9	US-10-087-192-1123	Sequence 1123, Ap
21	949	38.8	69090	15	US-10-203-138A-6819	Sequence 6819, Ap
22	948	38.8	567	14	US-10-203-138A-6819	Sequence 6819, Ap
23	863	35.3	473	14	US-10-203-138A-6819	Sequence 6819, Ap
24	833.5	34.1	57726	11	US-09-949-016-16430	Sequence 16430, A
25	833.5	34.1	73725	2	PCT-US02-41414-1208	Sequence 1208, Ap
26	813.5	33.3	72732	2	PCT-US02-41414-1205	Sequence 1205, Ap
27	654	26.8	533	10	US-09-513-999C-1340	Sequence 1340, Ap
28	654	26.8	533	12	US-09-513-999C-1340	Sequence 1340, Ap
29	650	26.6	420	14	US-10-203-138A-3751	Sequence 3751, Ap
30	650	26.6	454	14	US-10-203-138A-1690	Sequence 1690, Ap
31	511	20.9	1752	15	US-10-219-051B-13871	Sequence 13871, A
32	509.5	20.8	1876	15	US-10-170-235-24888	Sequence 24888, A
33	509.5	20.8	1876	18	US-60-452-680-295	Sequence 295, App
34	509.5	20.8	1901	15	US-10-219-051B-13873	Sequence 13873, A
35	499	20.4	1165	15	US-10-170-235-24404	Sequence 24404, A
36	498	20.4	318	18	US-09-513-999C-25950	Sequence 25950, App
37	498	20.4	318	10	US-09-513-999C-25950	Sequence 25950, A
38	498	20.4	318	12	US-10-170-235-34170	Sequence 34170, A
39	482	19.7	1581	15	US-60-455-444-673	Sequence 673, App
40	482	19.7	1581	19	US-60-455-444-673	Sequence 673, App
41	482	19.7	1581	14	US-60-455-444-673	Sequence 673, App
42	465	19.0	443	14	US-10-144-771-21735	Sequence 21735, A
43	460.5	18.8	2422	15	US-10-170-235-34449	Sequence 34449, A
44	460.5	18.8	2422	18	US-60-452-680-11304	Sequence 11304, A
45	460	18.8	1965	14	US-10-144-771-20068	Sequence 20068, A

ALIGNMENTS

RESULT 1
US-10-170-235-28132
Sequence 28132, Application US/10170235
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU
FILE REFERENCE: C1001380
CURRENT APPLICATION NUMBER: US/10/170, 235
CURRENT FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 42514
SEQ ID NO 28132
LENGTH: 1930
TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-28132

Alignment Scores: 2,47e-256 Length: 1930
Score: 2298.00 Matches: 426
Percent Similarity: 91.21% Conservative: 10
Best Local Similarity: 89.12% Mismatches: 14

Query Match: 94.03% Indels: 28
 DB: 15 Gaps: 2
 US-09-902-772-2 (1-451) x US-10-170-235-28132 (1-1930)

QY 2 AlaserthrilleysgluAlaLeuSerValValSerGluspgInserLeupheglucys 21
 DB 25 GCACCTCATATCAAGGAGGAGGCTTATCAGTGTGTGAGTACAGGACCGAGTGTGTGAGTGT 84
 QY 22 AlatyrglySerProHIsleuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
 DB 85 GCCTACGGAAGCCACACCTGGCTTAAGACAGATGACGCGTCTCTCTCCAGCGACTAT 144
 QY 42 GlyInThrSerIleuSerMetSerProArgValProGlnGlnAspIleuSerGlnProPro 61
 DB 145 GGACAGACTTCGAAGATGAGAGCCGCGTCCCTCAGCAGATGGCTGTCTCAACCCCA 204
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 DB 205 GCCAGGGTCACCATCAAAATGGAATGTAAACCTTACAGTGAATGGCTCAAGGACTCT 264
 QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGly 101
 DB 265 CCTGATGAATGCACTGTGGCCAAAGCGGGAAGATGGGGGAGCCAGACCCGAGACCGTGGG 324
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrAsn 121
 DB 325 ATGAACACAGGCGACACTCATGTGAGAGAGACATGCGACCCCAAAACATGACACAGAAC 384
 QY 122 GluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgGln 141
 DB 385 GAGGCGAGAGTTATCGCGCCAGCAGATCTACGCTATGAGTACAGACCATGTGCGGAG 444
 QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
 DB 445 TGGCTGAGTGGGGGGGGAAGATATAGCCCTTCACAGCGTCACATCTGTATTTCAG 504
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
 DB 505 AACATCATGTGGGAGGAGGAGTGTGCMAATGACCAAGACGACTCCAGAGGCTCACCCCC 564
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluArg----- 198
 DB 565 AGCTACACAGCGGCACATCTTCTCTCACATCTCCACTACAGAGACTCTCTTCCA 624
 QY 198 ----- 198
 DB 625 CATTGACTTCAGATGATGTGATAAAGCCTTACAAAACCTCCACGCTTAATGATGCT 684
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 214
 DB 685 AGAAACACAGGGGGTGCAGCTTTATTTCCCAATACCTCAGATATTCGGAAGTACG 744
 QY 215 GluArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgAspLeuArgLeuThr 234
 DB 745 CAAGAATTTCAACTAGGCCAGATTTACATATAGCCCCCAAGAGATCAGCTTGAGCC 804
 QY 235 SerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrValPro 253
 DB 805 GGTTCAGGCGCACCCACGCCCACTCGAAAGCTCTTAACCATCTCTTCCACAGTGGCC 864
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
 DB 865 AAAAAGTGAAGACCGCTCTCAGTTAGATCCTTATCAGATTTTGAGACCAAGATGAGC 924
 QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeuLeu 293
 DB 925 CGCCTTGCAATCCAGGCACTGGCCAGATCCAGTTTGGAGATTCCTCCGAGACTCTCG 984
 QY 294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMet 313
 DB 985 TCGGAGAGCTCCAACTCCAGCTGATCCTGGGAGAGCAACAGGGGAGTTCAAGATG 1044

QY 314 ThrAspProAspGluValAlaArgArgTrpGlyLysSerLysProAsnMetAsn 333
 DB 1045 ACGGATCCCGCAGGAGTGGGCCGCTGGGGAGAGCGGAAGACCAACCATGATGAAAC 1104
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
 DB 1105 TACGATAGCTCACCGCGGCTCCGTTACATATATACAAGAACATCATGACCAAGGTC 1164
 QY 354 HisLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
 DB 1165 CATGGGAAGCCCTACGCTTACAACTTGACTTCCACGGAGATGCGCCAGCCCTCCAGCC 1224
 QY 374 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
 DB 1225 CACCCCGGAGTATCTGTGTACAGTACCCCTCAGACCTCCGCTATGATGGCTCTAT 1284
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
 DB 1285 CACGCCACCCACAGAGATGAACTTGTGGCGCCACCCCTCCAGCCCTCCCGTGACA 1344
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 433
 DB 1345 TCTTCAGTTTCTTGTCTGCCCAAAACCCTACTGGAATTCACCACTGGGGGATATATAC 1404
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
 DB 1405 CCCAAGACTGAGGCTCCCGACAGCATATGCTTCTCATGTGGCGACTTACTAC 1458

RESULT 2
 US-60-452-680-11396
 : Sequence 11396, Application US/60452680
 : GENERAL INFORMATION:
 : APPLICANT: CARGILL, Michele
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 : FILE REFERENCE: CLO01450
 : CURRENT APPLICATION NUMBER: US/60/452,680
 : NUMBER OF SEQ ID NOS: 116213
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 11396
 : LENGTH: 1930
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-60-452-680-11396

Alignment Scores:
 Pred. No.: 2,47e-256 Length: 1930
 Score: 2298.00 Matches: 426
 Percent Similarity: 91.21% Conservative: 10
 Best Local Similarity: 89.12% Mismatches: 14
 Query Match: 94.03% Indels: 28
 DB: 18 Gaps: 2

US-09-902-772-2 (1-451) x US-60-452-680-11396 (1-1930)

QY 2 AlaserthrilleysgluAlaLeuSerValValSerGluspgInserLeupheglucys 21
 DB 25 GCACCTCATATCAAGGAGGAGGCTTATCAGTGTGTGAGTACAGGACCGAGTGTGTGAGTGT 84
 QY 22 AlatyrglySerProHIsleuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
 DB 85 GCCTACGGAAGCCACACCTGGCTTAAGACAGATGACGCGTCTCTCTCCAGCGACTAT 144
 QY 42 GlyInThrSerIleuSerMetSerProArgValProGlnGlnAspIleuSerGlnProPro 61
 DB 145 GGACAGACTTCGAAGATGAGAGCCGCGTCCCTCAGCAGATGGCTGTCTCAACCCCA 204
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 DB 205 GCCAGGGTCACCATCAAAATGGAATGTAAACCTTACAGTGAATGGCTCAAGGACTCT 264

QY 82 ProAspApcysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
 Db 265 CCGATGATGATGAGTGGCCAAAGGGGAGATGGTGGGAGCCAGACACCGGTTGGG 324
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrAsn 121
 Db 325 ATGAACATACGCGACGCTCATGGAGAGACACATGCGACCCCAACCAATGACACGAC 384
 QY 122 GluArgArgValIleValProAlaAspProThrIleuThrSerThrAspHisValArgGln 141
 Db 385 GAGCGCAGAGTTATCGGACGAGATCTACGCTAGAGTACAGACCATGTGGCGAG 444
 QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
 Db 445 TGGCTGAGTGGGGGGTGAAGAATATGGCTTCCAGATGTCACAACTCTGTTATTCAG 504
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
 Db 505 AACATCGATGGAGGAGGAGACGTGCAAGATGACCAAGACGACCTCCAGAGCTCACCC 564
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisIleuArgGluArg 198
 Db 565 AGCTACACACCCGACATCTTCTCTCATCTCCACTCAGAGAGACTCCCTTCCA 624
 QY 198 ----- 198
 Db 625 CATTTGACTTCAGATGATGTGATTAAGCCTTACAAACACTCCACGAGTTAATGATGCT 684
 QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 214
 Db 685 AGAACAACAGGGGCTGAGCTTTATTTTCCAAATTCCTGATATATCTCGAAGCTACG 744
 QY 215 GluArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThr 234
 Db 745 CAAGAATTCACACTAGGCCAGATTTACATATGAGCCCCCAGAGATAGCTTGAGCC 804
 QY 235 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 253
 Db 805 GGTACAGCGCCACCCAGCCGACGTCGAAAGCTGCTCAACCATCTCTCCACAGTCC 864
 QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
 Db 865 AAAACGTGAACACGAGCTCTCTAGTTAGATCTTATGATCTTGGACCAACAAGATAGC 924
 QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheIleuLeuGluLeu 293
 Db 925 CCGCTTGCATATCAGGACGATGGCCAGATCGCATGTCGCAATCTCTCGAGCTCTG 984
 QY 294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLysMet 313
 Db 985 TCGGACAGCTCCAACTCAGCTGATCAGCTGGGAAAGGCACCAAGGGAGTTCAAGATG 1044
 QY 314 ThrAspProAspGluValAlaArgArgTTPGlyGluArgLysSerLysProAsnMetAsn 333
 Db 1045 ACCGATGCCACAGAGTGGCCGCGCTGGGAGAGGGAAGCAACCAACCATGTAAC 1104
 QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 353
 Db 1105 TAGGATTAAGCTCAGCCGCCCTCCGTTACTACTATGACAAAGACATCAAGACAGAGTC 1164
 QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
 Db 1165 CAGGGAGAGCGGTACGCTCAAGTTGCACTCCAGGGATCGCCAGGCGCTCCAGCC 1224
 QY 374 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
 Db 1225 CACCCCGGAGATCTCTGTACAGTACCTCCAGACCTCCGTCACATGAGGCTCTAT 1284
 QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
 Db 1285 CAGGCCACCCACAGAGATGAACTTGTGGCGCCACCTCCAGGCGCTCCCGGTGACA 1344
 QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 433

Db 1345 TCTTCCAGTATTTTGTGCCCCCAACCACTAGTGAATTCACCACTGGGGGTATATAC 1404
 QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisIleuGlyThrTyr 451
 Db 1405 CCCAACACTAGGCTCCACCACAGCATATGCTTCATCTGCGGCACTTACTAC 1458
 RESULT 3
 US-10-170-235-28519
 ; Sequence 28519, Application US/10170235
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU
 ; FILE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
 ; FILE REFERENCE: C1001380
 ; CURRENT APPLICATION NUMBER: US/10/170, 235
 ; NUMBER OF SEQ ID NOS: 42514
 ; SEQ ID NO 28519
 ; LENGTH: 3294
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-10-170-235-28519
 Alignment Scores:
 Pred. No.: 3, 92e-244 Length: 3294
 Score: 2197.00 Matches: 410
 Percent Similarity: 91.96% Conservative: 13
 Best Local Similarity: 89.13% Mismatches: 23
 Query Match: 89.89% Indels: 14
 DB: Gaps: 4
 US-09-902-772-2 (1-451) x US-10-170-235-28519 (1-3294)
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 94 ATGGCCAGCACTATTAAAGAGCCCTTATCAGTTGTGAGTAGAGACCACTGTTGTTGAG 153
 QY 21 CysAlaTyrGlySerProHisIleuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 Db 154 TGTGCTTACGGAAGCCACCTGGCTTAACACAGATGACCGGCTCTCTCCAGCGAC 213
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 Db 214 TATGGAACAGACTTCCAAAGATGAGGCCGCGCTCCCTACGACGATTTGGTCTCAACCC 273
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 274 CCAAGCCAGGCTCACATCAAAATGGAATGTAAACCTTACCCAGGTGAATGGCTCAAGGAC 333
 QY 81 SerProAspApcysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 Db 334 TCTCTGATGAATGACATGTGGCCAAAGGGGAGATGGTGGGACCCAGACACCGTT 393
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
 Db 394 GGGATGAACATACGCGACGCTCATGAGAGAGAAGACATGCAACCCCAACATGACACG 453
 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuThrSerThrAspHisValArg 140
 Db 454 AACGAGCGCAGAGTTATCGTCCAGCAGATCTTACGCTATGAGTACAGACCAATGTGCGG 513
 QY 141 GluTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 Db 514 CAGTGGCTGAGAGGGGGGTGAAGAATATGCGCTTCCACACGTCACATCTGTATTC 573
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 574 CAGAACATCATGGGAAGAGAACTGTGCAAAATGACCAAGACATCTCCAGAGGCTCAC 633
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisIleuArgGluArgGlyAla 200
 Db 634 CCAAGTACAAACCCGACATCTTCTCATCTCCATCTCCATACCTCAGAGAGACTCTCTT 693

OY	201	ThrPheIlePheProAsnThrSerValIlyrProGluIaIarThrGlnIleThrThrArg	220
Db	694	CCACATTGTG-----ACTTCA-----GATAGTGTGATTAAGCCCTTACAAACTCT	738
OY	221	Pro-----AspleuProTyrgIugInAlaIarArgSerAla	232
Db	739	CCACGGTAAATGATGCTAGAAACAGATTTTACCATATGAGCCCCCAGAGATCAAGCC	798
OY	233	TrypThrSerHisSerHisProThr--GlnSerIysAlaThrGlnProSerSerThr	251
Db	799	TGGACCGGTGACGGCCACCCACCCCGCCAGTCGAAAGGTGCTCAACCATCTCTCTTCACA	858
OY	252	ValProIysThrGluAspGlnArgProGlnIleuAspProTyrgInIleLeuGlyProThr	271
Db	859	GTGGCCAAACATGAAGACCCAGCCGTCCTAGTAAATCCTATTCAGATTCTTGGACCAACA	918
OY	272	SerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglIleuLeuGlu	291
Db	919	AGTAGCCGCCCTTGGAAATTCAGGAGAGGCGCCAGATTCACGCTTTGGAGTTCTCTCTGGAG	978
OY	292	LeuIeuSerAspSerSerAsnSerAsnSerAlaIleThrTrpGluGlyThrAsnGlyGluPhe	311
Db	979	CTCTGTGTGGACAGCTCCCACTCCAGCTGCATCTCTGGAGACCAACGGGGAGATTC	1038
OY	312	LysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsn	331
Db	1039	AAGATGACGGATGCCGACGAGAGTGGCCCGCGCTGGGGAGAGCGAAGACCAACCCAC	1098
OY	332	MetAsnTyraAspIlysluSerArgAlaIleuArgTyrrTyraAspLysAsnIleMetThr	351
Db	1099	ATGAACTACGATAGCTACGCGCGCCCTCCGTTACTACTGACAAAGACATCATGACC	1158
OY	352	LysValHisGlyLysArgTyraIleTyrrLysPheAspPheHisGlyIleAlaGlnAlaIeu	371
Db	1159	AAGGTCCATGGGAAGCGCTACGCTACAAAGTTCACATTCACCGGAGATCGCCAGGCCCTC	1218
OY	372	GlnProHisProProGluSerSerMetTyrrLysTyrrProSerAspIeuProTyrrMetSer	391
Db	1219	CAGCCCCACCCCCCGAGGTCTCTGTACAAAGTACCCCTCAGACCTCCCTACATGGGC	1278
OY	392	SerTyrrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaIeuPro	411
Db	1279	TCCATATCAGCGCCACCCACAGAAATGATGACTTTGTGGCGCCACCTCCAGCCCTGCC	1338
OY	412	ValThrSerSerSerPhePheAlaIaIaProAsnProTyrrTrrpasSerProThrGlyGly	431
Db	1339	GTGACATCTTCCAGCTTTTGTTCGCCCAAAACCAATACGGAATTCACCAACTGGGGGT	1398
OY	432	IleTyrrProAsnThrArgLeuProAlaAlaHisMetProSerHisIleuGlyThrTyrr	451
Db	1399	ATATACCCCAACACTAGAGCTCCCAACCAAGCATATGCTTCTCATCTGGGCACTTACTAC	1458
RESULT 4			
US-60-452-680-11397			
Sequence 11397, Application US/60452680			
GENERAL INFORMATION:			
APPLICANT: CARGILL, Michele			
APPLICANT: GROUPE, Andrew			
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
FILE REFERENCE: C1001450			
CURRENT APPLICATION NUMBER: US/60/452,680			
NUMBER OF SEQ ID NOS: 116213			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 11397			
LENGTH: 3294			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-60-452-680-11397			
Alignment Scores:			

Coord. No.:	3,92e-244	Length:	3394
Score:	2197.00	Matches:	410
Percent Similarity:	91.96%	Conservative:	13
Best Local Similarity:	89.13%	Mismatches:	23
Query Match:	89.89%	Indels:	14
DB:	18	Gaps:	4

US-09-902-772-2 (1-451) x US-60-452-680-11397 (1-3294)

QY	1	20
QY	1 Metalasertmrlllelyglualaleuservalvalsergluasplinsertleuphegu	20
Db	94 ATGGCCACACTATTATAGGAAGCCCTATACAGTTGTGAGTAGAGACAGACGCTGTGTGAG	153
QY	21 CysalatrlyglsyerProHisteulalysrtrnglmetrthlaSersersergergu	40
Db	154 TGTGGCTACGGNACCCACACCTGGCTTAAGACAGATGACCGCTCTCTCCACGGAC	213
QY	41 TyrglylntrhSerLysmetsSerProarqvalProglnglnasPTPLeuserlpro	60
Db	214 TATGACAGACTTCCAGATGAGCCACGCGCTCCCTCAGACAGATGTGGCTGCTCAACC	273
QY	61 ProAlaargvalThrlllelysmetlucysasnProasnglnvalasnglsyerargsn	80
Db	274 CCAGCCAGGCGCACCATCAATAATGAATGAATCCTACGACAGTAAATGGCTCAAGAAC	333
QY	81 SerProaspaPcySerSerValalalysgllylvsmetvalSerSersersaspaVal	100
Db	334 TCTCCTGATGATGCAGTGTGGCCCAAGCGGGAGAGATGTGTGGACGCCAGACACCGTT	393
QY	101 GlymetasnyrglySerTyrmelglulysnshisileproProasnmethrThr	120
Db	394 GGGATGACACTAGCGCAGCTACATGAGAGAAAGACATGCCACCCCAACATGACACAG	453
QY	121 AsnglualargvalIlevalProalaspProthrleutpserThrAsphistalarg	140
Db	454 AACGGAGCCCAAGTTATGTGGCCACCAATCCTACGATATGAGATGACAGACCATGTGGG	513
QY	141 GlntPleauglUTPrAlaVallylsylunrglyleuProaspaValAsplleleuPhe	160
Db	514 CAGTGGCTGGAGTGGGGCGGTGAAGAATATGGCTTCCAGATGTCACATCTTGTATTTC	573
QY	161 GlnasnllleaspglyLysgluleucysLysmethrLysaspaPheglnaargleuthr	180
Db	574 CAGAAATCATGTGGGAAGAACTGTGCAAGATGACCAAGAGACGACTTCCAGAGCTCACCC	633
QY	181 ProserTyranAlaAsplleleuenseerHisteuHslyrleuAlarglualarglAla	200
Db	634 CCAGACTCAACAGCCGACATCTTCTCCACTTCCACTCCACTCAGAGAGCTCTCTCT	693
QY	201 ThrPheillePheProAsnThrSerValtyrProglualathrglnAlarglIlethrthArg	220
Db	694 CCACATTTG-----ACTTCA-----GATGATGTGTGAATMAAGCCTTCAAAACCT	738
QY	221 Pro-----AspleuProTyrglulAlalargArgserAla	232
Db	739 CCACGGTAAATGCAATGCTAGAAACACACATTTACATATGAGCCCCACAGGAGATCACCC	798
QY	233 TrpThrSerHsSerHsProThr---GlnserLysAlaThrGlnProserSerserThr	251
Db	799 TGGACCGGTGACGGCCACCCACGCCCCAGTGTGAAAGGTGTCAACCACTTCCTTCCACA	858
QY	252 ValProlystrnrgluaspglnArgProglneuaspproTyrglnlleleuglyProthr	271
Db	859 GTGCCCAAAACTGAAGACACAGCTCTCAGTTAGATCTTACAGATTCTTGGACCAACA	918
QY	272 SerSerArgleuAlaasnProglySerlyGlnlleGlnleuTyrglnPheleuLeuglu	291
Db	919 AGTAGCGCCCTTGCAAAATCCAGGAGAGTGGCCAGATCCAGCTKGTGGCACTTCTCTGGAG	978
QY	292 LeuLeuSeraspSerSerAsnSerAsnscylIlethrTrpGluGlyThrAsnglygluPhe	311
Db	979 CTCTGTGTGGACACTCCAACTCCAGCTGCATCACCTGGGAAGGACCAACGGGGAGTTTC	1038

```

Oy 312 LysMetThrAspProaspGluValAlaIaArgArgTrpGluArgLysSerLysProasn 331
Db 1039 AAGATGACGAGATCCCGAGAGAGTGGCCGGCGCTGGGAGAGCGGAAAGCAACCCAC 1098
Oy 332 MetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThr 351
Db 1099 ATGAACCTACGATTAACCTACAGCGCGCCCTCCGTTACTACTATGACAAACATCATGTGACC 1158
Oy 352 LysValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGluAlaLeu 371
Db 1159 AAGGTCATGGGAACCGCTAGCGCCACAGATTCGACTTCACGGGATGCCCGAGCCCTC 1218
Oy 372 GluProHisProProGluSerSerMetLysTyrTyrProSerAspLeuProTyrMetSer 391
Db 1219 CAGCCCGACCCCGCGAGTACTCTGTACAAAGTACCCCTCAGACCTCCCGTACATGGGC 1278
Oy 392 SerTyrHisAlaHisProGluLysMetLysPheValAlaIaProHisProProAlaLeuPro 411
Db 1279 TCCATATCAGCCCGCCACAGAAAGATGTAACCTTGTGGCGCCCGCCACCTCCAGCCCTCC 1338
Oy 412 ValThrSerSerSerPhePheAlaIaIaProAsnProTyrTrpAsnSerProThrGlyGly 431
Db 1339 GTGACATCTTCAGATTTTTCCTGCTGCCCCAACCATACGTGAATTCACCAACTGGGGGT 1398
Oy 432 IleTyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1399 ATATATCCCAACACTAGGCTCCCAACGACATATGCTTCTCATCTGGGACCTTACTAC 1458

RESULT 5
PCT-US02-04915-41
; Sequence 41, Application PC/TUS0204915
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Susan R.
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200PC
; CURRENT APPLICATION NUMBER: PCT/US02/04915
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/285,475
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-04915-41

Alignment Scores:
Pred. No.: 4,66e-243 Length: 1389
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 2 Gaps: 4

US-09-902-772-2 (1-451) x PCT-US02-04915-41 (1-1389)
Oy 2 AlaSerThrIleLysGluAlaLeuSerValAlaSerGluAspGlnSerIleuPheGluCys 21

```

```

Db 25 GCAGCTCATATCAAGGAAGCCCTATACATCTGTGATGAGACACAGTCTTTTGTAGT 84
Oy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41
Db 85 GCCTACGGAGACGACACCTGTGTAAAGACAGATGACAGGGGTCTCTCTCCAGGACATAT 144
Oy 42 GlyAlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGluProPro 61
Db 145 GGACAGACTTCACAGATAGGCCCGGCTCCCTCAGCAGGATGGCTCTTCACACCCCA 204
Oy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGGCTCACCATCAATAATGAAATGAACCTACACCGAGTGAATGGCTCAAGAACTCT 264
Oy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 265 CCTGATGAATACAGTGTGGCCAAAGCGGGAAGATGGTGGGACGCCAGACCGTTTGG 324
Oy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrAsn 121
Db 325 ATGAACTACGGCAGCTACATGAGAGAGACATGCCCCCAACATACACAGAC 384
Oy 122 GluArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArgGln 141
Db 385 GAGCGCAGAGATTATCGTCCAGCAGATCTCTACGGCATGAGATACAGACATGTGGCGAG 444
Oy 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161
Db 445 TGGCTGAGATGGCGGTGAAGAATATGGCTTCAGACGTCACATCTGTATTCAG 504
Oy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCGATGGGAAGAACTGTGCAAGTGCACAAAGGACGACTTCACAGGCTCACCC 564
Oy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db 565 AGCTACAAAGCGCAGCATCTCTCTCATATCTCCACATCTCCAGAGACATCCCTTCCA 624
Oy 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrArgPro 221
Db 625 CATTTG-----ACTTCA-----GATATGTGATAAAGCTTCAAAACTCTCA 669
Oy 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTrp 233
Db 670 CGGTTATGCTGCTAGAAACACAGATTTACCATATGAGCCCGCCAGAGATCAGCTGG 729
Oy 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 252
Db 730 ACCGCTACGGCCACCCCGCCAGCGCCAGTCAAGCAACATCTCTCCACAGTGT 789
Oy 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db 790 CCAAAACTGAAGACCGCGCTCAGTATGATCTTATCATGATTTCTGGACCAACAGT 849
Oy 273 SerArgLeuAlaAsnProGluSerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeu 292
Db 850 AGCGGCTGCAAAATCCAGGCGAGTGGCCAGATCCAGCTTGGCAGTTCTCTGAGCTC 909
Oy 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlnIleThrAsnGlyLysPheLys 312
Db 910 CTGTGGGACAGCTCCAACTCCAGCTGATACCTGGGAGGACCAACCGGGGATTTCAAG 969
Oy 313 MetThrAspProaspGluValAlaIaArgArgTrpGluArgLysSerLysProAsnMet 332
Db 970 ATGACGAGATCCCGAGCGAGTGGCCCGCGTGGGAGAGCGCGAAGACCAACCAACATG 1029
Oy 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db 1030 AACTACATTAAGCTCAGCGCGCTCGTTACTACTATGACAAACATCATATACCAAG 1089
Oy 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372

```

```

Db      1090 GTCATGGAGCGCTACGCTTACAGTTCACACGGATGCCAGGCCCTCCAG 1149
Qy      373 ProHisProProgluSerMetTyrlYsTyProSerAspLeuProTyrmMetSer 392
      1150 CCCACCCCCGGGATCTCTGTACAGAACCCCTCAGACCTCCGCTACATGGGCTCC 1209
Qy      393 TyHHisAlaHisProGlnLysMetAsnPhenValAlaProHisProProAlaLeuProVal 412
      1210 TATCACGCCACCCACAGAGATGAACTTGTGGCCGCCACCTCCAGCCCTCCCGCTG 1269
Db      413 ThisSerSerPhePheAlaAlaProAsnProTyTrpAsnSerProThnglylyle 432
      1270 ACATCTTCCAGTTTGTGCTGCCAACCCATCTGGAATTCACCACTGGGGGTATA 1329
Qy      433 TyProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuglyltyrtyr 451
      1330 TACCCCAACACTAGGCTCCCAACACGCAATATGCTTCTCATCTGGGCACTTACTAC 1386

RESULT 6
US-10-211-462-41
: Sequence 41, Application US/10211462
: GENERAL INFORMATION:
: APPLICANT: Murray, Richard
: APPLICANT: Glynn, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Aziz, Nafana
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
: FILE REFERENCE: 018501-006200US
: CURRENT APPLICATION NUMBER: US/10/211,462
: CURRENT FILING DATE: 2003-02-13
: PRIOR APPLICATION NUMBER: US 09/784,356
: PRIOR FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: US 09/791,390
: PRIOR FILING DATE: 2001-02-22
: PRIOR APPLICATION NUMBER: US 60/310,025
: PRIOR FILING DATE: 2001-08-03
: PRIOR APPLICATION NUMBER: US 60/334,244
: PRIOR FILING DATE: 2001-11-29
: NUMBER OF SEQ ID NOS: 230
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 41
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-211-462-41

Alignment Scores:
Pred. No.: 4.66e-243 Length: 1389
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 89.67% Mismatches: 24
Query Match: 89.28% Indels: 14
DB: 14 Gaps: 4

US-09-902-772-2 (1-451) x US-10-211-462-41 (1-1389)
Qy      2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
      25 GCAGCCTCATATCAAGAGCGCTTATCATGCTGTGAGTAGAGACCACTGCTTGTAGTGT 84
Db      22 AlaTyrlGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTy 41
      85 GCCTAGCGAAGCGCACACCTGCTAGACAGAGATGACCGCTCTCTCCACAGCACTAT 144
Qy      42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
      145 GGACAGACTTCCAGATGAGACCCACCGCTCCACAGAGAGATTGGCTGTCAACCCCA 204
Db      62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
      205 GCCAGGGTCAACCAAAATGGAATGATTAACCTTAGCCAGGTGAATGGCTCAAGAACTCT 264

```

```

Qy      82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
      265 CCGTATGAATGACAGTGTGGCCAAAGCGGGAAGATGTTGGAGACCCAGACACCGTTGGG 324
Db      102 MetAsnTyrlGlySerTyrmMetGluGluHisIleProProProAsnMetThrAsn 121
      325 ATGAACATACGGCAGCTCATGAGAGAACACACATGCGACCCCAACATGACACAGAAC 384
Qy      122 GluArgArgValIleValProAlaAspProThrLeuProSerThrAspHisValArgGln 141
      385 GAGCGCAGAGTTATGTCGCCAGCATCTTACGCTATGAGATGACAGCACTGTGGCCAG 444
Db      142 TrpLeuGluTrpAlaValLysGluTyrlGlyLeuProAspValAspIleLeuLeuPheGln 161
      445 TGGCTGAGTGGGGGGAAGAAATATGCGCTTCACAGCTCAACATCTGTATTTCAG 504
Qy      162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
      505 AACATCGATGGAGGAACTGTGCMAATGATACCAAGAGACATTCACAGAGCTCACCCCC 564
Db      182 SerTyraAsnAlaAspIleLeuLeuSerHisLeuHisIleTyrlLeuArgGluArgValAlaThr 201
      565 AGCTACACCGCCGACATCTTCTCTCATCTCCACACTCCACAGAGACTCCTCTCCA 624
Qy      202 PheIlePheProAsnThrSerValTyrlProGluAlaThrGlnArgIleThrArgPro 221
      625 CATTTG-----ACTTCA-----GATGATGTTGATTAACCTTCAAAACTCTCCA 669
Db      222 -----AspLeuProTyrlGluGlnAlaArgAspSerAlaTrp 233
      670 CGGTTAATGCATGCTAGAAACACAGATTAACCATATGAGCCCCAGAGATGACCTGG 729
Qy      234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 252
      730 ACGGTCAGGCGCACCCACGCGCCAGTGAAGAGCTGCACCATCTCTTCCACAGCTG 789
Db      253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrlGlnIleLeuGlyProThrSer 272
      790 CCCCAAACTGAAGCCAGCGCTCTCATGATGATTCATTATGATTCAGTTCGACCAACAAAGT 849
Qy      273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuArgPheLeuLeuGluLeu 292
      850 AGCGGCTTCAAAATCCAGGCAGTGGCCAAATCCAGCTTGGGAGTTCCTCGGAGCTC 909
Db      910 CTGTCGACACGCTCCAACTCAGCTGCATCAGCTGGAGAGCAACCAAGGGAGTTCAAG 969
Qy      293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLupheLys 312
      910 CTGTCGACACGCTCCAACTCAGCTGCATCAGCTGGAGAGCAACCAAGGGAGTTCAAG 969
Db      313 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
      970 ATGACGGATCCCGACAGGAGTGGCGCGCTGGGGAGAGCGGAAGCAAAACCAACATG 1029
Qy      333 AsnTyraAspLysLeuSerArgAlaLeuArgTyrlTyrlAspLysAsnIleMetThrLys 352
      1030 AACTAGATTAAGCTCAGCGCGCTCCGCTGTTACTACTATGACAAAGACATCATGACCAAG 1089
Db      353 ValHisGlyLysArgTyrlAlaTyrlLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
      1090 GTCATGGAGCGCTACAGCGCTTACAGTTCACAGGAGATGCGCCAGCGCTCCAG 1149
Qy      373 ProHisProProgluSerMetTyrlYsTyProSerAspLeuProTyrmMetSer 392
      1150 CCCACCCCCGGGATCTCTGTACAGAACCCCTCAGACCTCCGCTACATGGGCTCC 1209
Db      393 TyHHisAlaHisProGlnLysMetAsnPhenValAlaProHisProProAlaLeuProVal 412
      1210 TATCACGCCACCCACAGAGATGAACTTGTGGCCGCCACCTCCAGCCCTCCCGCTG 1269
Qy      413 ThisSerSerPhePheAlaAlaProAsnProTyTrpAsnSerProThnglylyle 432
      1270 ACATCTTCCAGTTTGTGCTGCCAACCCATCTGGAATTCACCACTGGGGGTATA 1329

```

OY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db 1330 TACCCCAACACTAGGCTCCCAACAGCATATGCTTCTCATCTGGGACCTTACTAC 1386

RESULT 7
US-10-087-192-1127
; Sequence 1127, Application US/10087192
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1127
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1127

Alignment Scores:
Pred. No.: 4,66e-243 Length: 1389
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 88.67% Mismatches: 24
Query Match: 89.28% Indels: 14
Gaps: 4

US-09-902-772-2 (1-451) x US-10-087-192-1127 (1-1389)

OY 2 AlaSerThrIleuLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db 25 GCAGCTCATATCAAGAGAAAGCTTATCACTGTGAGTGGACACACTGTTGTTGAGTGT 84

OY 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
Db 85 GCCTACGAGAACGCCACCTGCTGCTAAGACAGAGATGACGGGTCTCTCCACGAGACAT 144

OY 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnProPro 61
Db 145 GGACAGACTTCCAGATGAGCCACGCGTCCCTCAGCAGGATGGCTGTCTCAACCCCA 204

OY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGGGTCAACATCAAAATGGAATGTAACCTTACGCCAGGTGAATGGCTCAAGGAACCT 264

OY 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101
Db 265 CCTGATGATGCAGATGTCGCAAAAGCGGGAAGATGGTGGGACAGCCAGACACCGTTGG 324

OY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
Db 325 ATGACATACGGCAGCTACATGAGAGAGACATGACACCCCAAAACATACACACGAMC 384

OY 122 GluArgValIleIleValProAlaAspProThrLeuThrPheSerThrAspHisValArgGln 141
Db 385 GAGCGCAGAGATTATCTGTCACGACAGATCTTACGCTATGAGATACAGACATGTGGGAG 444

OY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 445 TGGCTGAGATGGCGGTGAAGAATATGCTTCACAGCGTCAACATCTTTATTTCAG 504

OY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCATGGGAAGAACTGTGCAGATGACCAAGGACGACTTCCAGAGGCTCACACCCC 564

OY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db 565 AGCTACAAAGCGGACATCTCTTCTACATCTCCACTACCTACAGAGACATCTCTTCCA 624

OY 202 PheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThrArgPro 221
Db 625 CATTTG-----ACTTCA-----CATGATGTTGATTAAGCCTTCAAAACATCTCCA 669

OY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTrp 233
Db 670 CGGTTAATGATGCTAGAAACACACAGATTATACCATTAAGAGAGAGATACAGCTTGG 729

OY 234 ThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db 730 ACCGTCACGGCCACCCACAGCCCGCCAGTGAAGCTGCTCAACATCTCTTCCACAGTG 789

OY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db 790 CCAAAACTGAAGACACAGCTCTCAGTTAGATCTTATTCAGATCTTGGACCAACAGT 849

OY 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheGlnPheLeuLeuLeu 292
Db 850 AGCGCCTTGCATAATCCAGCAGATGCGCAGATCCAGCTTGGCAGTCTCTCCGAGGCTC 909

OY 293 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLys 312
Db 910 CTGTGGACAGCTCAACTCCAGCTGCAATCAGCTGGAAGAGGACCAAGGGGAGTTCAG 969

OY 313 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
Db 970 ATGACAGATCCCGACAGAGAGTGGCCCGGCGTGGGAGAGCGGAAGACCAACCAATG 1029

OY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db 1030 AACTAGATTAAGCTCAGCGCGCCCTCGTTACTATGACAAACATCATGACCAAG 1089

OY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
Db 1090 GTCCATGGGAAGCGTACCGCTACAGTTCGACTTCCACGGGATGCCAGGCCCTCCAG 1149

OY 373 ProHisProGlnSerSerMetLysTyrTyrProSerAspLeuProTyrMetSerSer 392
Db 1150 CCCCACCCCGGAGATCATCTGTGACAGTAAGTACCCTCAGACCTCCCTACATGGGCTCC 1209

OY 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db 1210 TATCAGCGCCACCAACAAATGAATTTGTGGGCCCCCAGCCCTCCCGCTG 1269

OY 413 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyTyrIle 432
Db 1270 ACATCTTCCAGTTTCTTGTGCTGCCCCAACCATTAAGTGAATTCACCAACTGGGGTATTA 1329

OY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
Db 1330 TACCCCAACACTAGGCTCCCAACAGCATATGCTTCTCATCTGGGACCTTACTAC 1386

RESULT 8
US-10-126-052A-330
; Sequence 330, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Richard
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
; FILE REFERENCE: 018501-001530US
; CURRENT APPLICATION NUMBER: US/10/126,052A
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/339,245

```

: PRIOR FILING DATE: 2001-11-09
: PRIOR APPLICATION NUMBER: US 60/350,666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/334,370
: PRIOR FILING DATE: 2001-11-29
: PRIOR APPLICATION NUMBER: US 60/372,246
: PRIOR FILING DATE: 2002-04-12
: NUMBER OF SEQ ID NOS: 691
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 330
: LENGTH: 3166
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-126-052A-330

Alignment Scores:
Pred. No.: 2,05e-242 Length: 3166
Score: 2182.00 Matches: 407
Percent Similarity: 91.72% Conservative: 14
Best Local Similarity: 86.67% Mismatches: 24
Query Match: 89.28% Indels: 14
Gaps: 4
US-09-902-772-2 (1-451) x US-10-126-052A-330 (1-3166)

```

```

OY 2 AlaserthrlielysgluAlaleuSerValValsergluaspInserleupheglucys 21
Db 281 GCACCTCATATCAAGAGAGCCCTTATCAGTGTGAGTGAGCAGCAGCTGTTGAGTGT 340
OY 22 AlatyrglyserProHisleuAlaleuSerThrleuMetThrAlaserSerSerSerSglTyr 41
Db 341 GCCTACGAGACGCCACACCTGCTGCTAGACAGATGACGCGCTCTCTCCAGGACATAT 400
OY 42 GlylnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnProPro 61
Db 401 GGACAGACTTCCAGATGAGCCACGGCTCCCTCAGCAGATGGCTGTCTCAACCCCA 460
OY 62 AlaryValThrleuLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 461 GCCGGGTCCACCAATGGAATGTAACTTACGAGGAGGTAAGTGGCTCAAGCAACTCT 520
OY 82 ProAspCysSerValAlaleuGlyLysMetValSerSerSerSerAspAsnValGly 101
Db 521 CCTATGATGACAGTGTGGCCAAAGCGGGAAGTGGTGGCAGCCAGACCGGTGG 580
OY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThrAsn 121
Db 581 ATGAACATGACGACATGATGAGAGAGACATGACACCCCAACATGACACGAC 640
OY 122 GluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgGln 141
Db 641 GAGGCGAGAGTTATCGCCAGCAGATCTTACGCTATGAGTACAGACCATGTCCGCGAG 700
OY 142 TyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 701 TGGCTGAGTGGCGGGGGAAGATATGCGCTTCAACGCAACATCTTGTATTCAG 760
OY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATCATGGGAAGGAAGATGTGCAAGATGACCAAGACGACTCCAGAGGCTCACCC 820
OY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThr 201
Db 821 AGCTACACGCCGACATCTTCTCTCACATCTCCACTACAGAGAGACTCCTTCCA 880
OY 202 PheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgPro 221
Db 881 CATTTG-----ACTTCA-----GATGATGTTGATTAAGCCTTACAAACTCTCA 925
OY 222 -----AspLeuProTyrGluGlnAlaArgArgSerAlaTyr 233
Db 926 CGGTTATGATGCTAGAACACAGATTTTACCATATGAGGCCGCCAGAGATTCAGCCTGG 985

```

```

OY 234 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 252
Db 986 ACCGGTCACGGCCGCCACCCACCCAGTGGGAAGCTGTCAACCATCTCTTCCACAGTG 1045
OY 253 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
Db 1046 CCCAAACTGAAAGCACAGCTGCTCAGTTACATCTTATGATTTCTTGACCAACAGT 1105
OY 273 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGluLeu 292
Db 1106 AGCCGCTTGCAAAATCCAGGAGTGGCCAGATCCAGCTTTGGAGTTCCTCTGAGCTC 1165
OY 293 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLys 312
Db 1166 CTGTGCGACAGCTCACTCAGTCCAGTCACTCAGGGAAGCACCACAGGGGAAGTTCAG 1225
OY 313 MethrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 332
Db 1226 ATGACGATCCCGACGAGGTGGCCGCGCTGGGGAGAGCGGAAGACCAACCCACATG 1285
OY 333 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 352
Db 1286 AACTACGATAGCTCAGCGCGCCCTCGTACTACTATGACAAAGACATCATGACCAAG 1345
OY 353 ValHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 372
Db 1346 GTCCATGGGAAGCCCTCAGCTTACAGTTCACCTTCCACGCGGATGCCAGCCCTCCAG 1405
OY 373 ProHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 392
Db 1406 CCCACCCCGCGAGTACTCTGTACAGTACCCCTCAGCCTCCGTTATGAGCTCC 1465
OY 393 TyrHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 412
Db 1466 TATACAGGCCACCCACAGAGATGACTTGTGGCGCCACCCCTCAGCCCTCCCGCTG 1525
OY 413 ThrSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLys 432
Db 1526 ACATCTTCCAGTCTTTTGTGCGCCCAACCCATACGGAATTCACCAACTGGGGGTATA 1585
OY 433 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 451
Db 1586 TACCCCAACTAGGCTCCACCAAGCCATATGCTTCTCATCTGGGCATTTACTAC 1642

RESULT 9
US-10-087-192-1124
: Sequence 1124, Application us/10087192
: GENERAL INFORMATION:
: APPLICANT: Morris, David W.
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
: FILE REFERENCE: 529452000122
: CURRENT APPLICATION NUMBER: US 10/087,192
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 09/747,377
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/798,586
: NUMBER OF SEQ ID NOS: 2059
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1124
: LENGTH: 1297
: TYPE: DNA
: ORGANISM: Mus musculus
US-10-087-192-1124

Alignment Scores:
Pred. No.: 1.57e-228 Length: 1297
Score: 2057.00 Matches: 384
Percent Similarity: 83.69% Conservative: 11
Best Local Similarity: 81.36% Mismatches: 9
Query Match: 84.17% Indels: 68

```

DB: 15 Caps: 2
 US-09-902-772-2 (1-451) x US-10-087-192-1124 (1-1297)

QY 7 GluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrGlySerPro 26
 |||||
 Db 2 GAGGCGCTTGTGACGTGTGACGAGGAGCAGTCACTATTGTAGTGTGCTACGGAGCCCA 61
 |||||

QY 27 HisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyGlnThrSerLys 46
 |||||
 Db 62 CACCTGGCTTAAGACAGATGACCGCATCTCTCCAGTACATGAGCCAGACATCCAG 121
 |||||

QY 47 MetSerProArgValProGlnGlnAspThrLeuSerGlnProAlaArgValThrIle 66
 |||||
 Db 122 ATGAGTCCAGAGTCCCTCCAGAGAGTGGCTGTCTCAAGCCCGACGAGGTCACCATC 181
 |||||

QY 67 LysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspCysSer 86
 |||||
 Db 182 AAGATGGAGTGAACCTAGTCAAGTGAATGTTCCAGAACTCAGCTGATGATGAGT 241
 |||||

QY 87 ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySer 106
 |||||
 Db 242 GTAAACAAAGGTGGGAAGATGTGGGCGACCGGATACGTGGGGATGAGAGCTACGGCAGC 301
 |||||

QY 107 TyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluArgArgValIle 126
 |||||
 Db 302 TACATGAGAGGAGAAGCATGTGCCGCTCCCAATATATGACCAATATGAGCCAGAGTATC 361
 |||||

QY 127 ValProAlaAspProThrLeuThrPheThrAspHisValArgGlnThrLeuGluTyrPala 146
 |||||
 Db 362 GTCCCTGCAGATCCTACTGTGTGAGACAGACCAATGTCGAGACGTGGAGTGGCGC 421
 |||||

QY 147 ValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166
 |||||
 Db 422 GTAAAGAAATATGCGCTCTCGATGTGAGCGTCTTACTATTGAGATATCGATGGGAG 481
 |||||

QY 167 GluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186
 |||||
 Db 482 GAGCTGTGCAAGATGACAAAGATGACTTCCAGCGGTCTCAGCGGAGTCAATGCGGAC 541
 |||||

QY 187 IleLeuLeuSerHisIleuHisTyrLeuArgGluArg 198
 |||||
 Db 542 ATCTCTCTCCACATCTCCACTACACAGAGACATCCCTCCACATCTGACTCCGAT 601
 |||||

QY 199 199 Gly 199
 |||||

Db 602 GACGTGATAGGCTTTACAAAACCTCCACGGTTAATGATGCCAGAAACAGAGGGGT 661
 |||||

QY 200 AlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThr 219
 |||||
 Db 662 GCAGCTTTATTTTCCCAATACTTCAATATATCCGAAAGCTTACGAAAGATATACACT 721
 |||||

QY 220 ArgProAspLeuProTyrGluGlnAlaArgArgSerAlaThrThrSerHisSerHisPro 239
 |||||
 Db 722 AGGCA 727
 |||||

QY 240 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArg 259
 |||||
 Db 727 727
 |||||

QY 260 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGly 279
 |||||
 Db 728 GATCTTACCGATCCGATCCGAGCGGAGCGACAGTACCGCTGTGTAATCCAGGT 778
 |||||

QY 280 SerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeuSerAspSerAsnSer 299
 |||||
 Db 779 AGTGGCCAGATCCAGCTGTGGCGAGTTCCTGCAACTCCTGTGCAGACAGCTCCAACTCC 838
 |||||

QY 300 AsnCysIleThrThrGluThrThrAsnGlyGluPheLysMetThrThrAspProAspGluVal 319
 |||||
 Db 839 AACTGCATCACCTGGGAGGACCAAGGAGGAGTGAAGTGAACAACCCGAGGAGGTG 898
 |||||

QY 320 AlaArgArgTyrPylGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 339
 |||||

Db 899 GCTCGGCGCTGGGGGAGAGAGAGACACCCACATGAACTATGACAAAGCTACGCGC 958
 |||||

QY 340 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisGlyLysArgTyrAla 359
 |||||

Db 959 GCCCTCCGCTACTACTACGACAAACATCATGACCAAGGTGACGAGGAGCCCTACGCC 1018
 |||||

QY 360 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProProGlnSerSer 379
 |||||

Db 1019 TACAACTTACTTCCACAGGGAATGGCCAGGCGCTGCAGCCACCTCTGAGTCGTC 1078
 |||||

QY 380 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisAlaHisProGlnLys 399
 |||||

Db 1079 CTGTACAGATACCCCTCCGACCTGCCATATGAGGCTCTATACGCCACCCAGAG 1138
 |||||

QY 400 MetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSerPhePheAla 419
 |||||

Db 1139 ATGAAGTTTGTGTCCCCACCTCCGCTCTCCAGTCAATCTTCCAGTCTTGTGT 1198
 |||||

QY 420 AlaProAsnProTyrThrPasnSerProThrGlyGlyIleTyrProAsnThrArgLeuPro 439
 |||||

Db 1199 TCCCGAACCCTACTGTGAATTCACGAGTGGGGGCTATCAACCGAAGACTAGGCTCCA 1258
 |||||

QY 440 AlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451
 |||||

Db 1259 GCCAGCATATGCCCTCTCTACCTGCGCAGCTACTAC 1294
 |||||

RESULT 10
 US-10-144-771-355
 ; Sequence 355, Application US/10144771
 ; GENERAL INFORMATION:

; APPLICANT: VENTUR, J. Craig
 ; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
 ; FILE REFERENCE: CL001321
 ; CURRENT APPLICATION NUMBER: US/10/144, 771
 ; CURRENT FILING DATE: 2002-05-15
 ; NUMBER OF SEQ ID NOS: 47235
 ; SEQ ID NO 355
 ; LENGTH: 2172
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-10-144-771-355

Alignment Scores:

Pred. No.: 1,99e-220 Length: 2172
 Score: 1991.00 Matches: 374
 Percent Similarity: 80.54% Conservative: 11
 Best Local Similarity: 78.24% Mismatches: 9
 Query Match: 81.46% Indels: 84
 DB: 14 Gaps: 2

US-09-902-772-2 (1-451) x US-10-144-771-355 (1-2172)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 |||||

Db 107 ATGGCAGACGACCTAATAAGAGGCGCTGTGACGTTGTGAGCGAGAGACAGTACATTTGAG 166
 |||||

QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGln 40
 |||||

Db 167 TGTGCTACAGGAGCGCACACCTGCTGAAGACAGATGACCCGATCTCTCCAGAGAC 226
 |||||

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60
 |||||

Db 227 TATGGCCAGACATCCAAAGATGAGTCCAGAGTCCCTCAGGAGAGTGGCTGTCAAGCC 286
 |||||

QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 |||||

Db 287 CCAGCCAGGCTCCACATCAAGATGAGTGAACCTAGTCAAGTGAATGTTCCAGAGAC 346
 |||||

QY 81 SerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnVal 100
 |||||

Db 347 TCACCTGATGACGTGAGTGTGAACAAGGTGGGAGAAATGATGTGGCAGCCCGATCTACTGTG 406
 |||||

QY 101 GlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThr 120
Db 407 GGGATGCTACCGCAGCTACATGAGAGAGAGATGCGCGCTCCCATATATGACACA 466
QY 121 AsnGluArgValIleValProAlaAspProThrLeuThrPheSerThrAspHisValArg 140
Db 467 AATGAGCGCAGAGTGCATGCTCCGAGATCTACTGTGGACACAGACCATGTCGA 526
QY 141 GlnThrLeuGluTrpAlaValIleGlyGluTyrGlyLeuProAspValAspIleLeuPhe 160
Db 527 CAGTGGCTGAGTGGCGGGTGAAGAAGATATGGCGCTCCGATGTGGAGCGTCTACTATT 586
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 587 CAGAAATTCATGGAGAGAGCTGTGCAAAATGCAAAAGATGACTTCCAGCGGCTCAGC 646
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgValArg 198
Db 647 CCGAGCTACAAATGCCGACATTTCTCTCATCTCCACTACCTCAGAGAGATCCCTT 706
QY 198 ----- 198
Db 707 CCACATCTGACTTCCGATGAGCTTGATAGGCTTTACAAAACCTCCAGCGTTAATGAT 766
QY 199 -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 213
Db 767 GCCAGAAACACAGGGGCTGCAGCTTTATTTTCCAAATATCTCAGATATCCGAGACT 826
QY 214 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 233
Db 827 ACCGAAAGAAATTCACCTAGG----- 847
QY 234 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValPro 253
Db 847 ----- 847
QY 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
Db 847 ----- 847
QY 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 293
Db 848 -----CCAGTATGCGCAGATCCAGCTGTGGCAGTTCTCTCGTGCAGCTCTG 895
QY 294 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysMet 313
Db 896 TCAGACAGCTCCAACTCACTCACTCACTGCGAAGGACCAAGGGGAGTTCAATG 955
QY 314 ThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsn 333
Db 956 ACAGACCCGGAGCGAGTGGCTGCGCGGGGAGAGAGCAAGCAAGCCCAACATGATAC 1015
QY 334 TyrAspLysLeuSerArgAlaLeuArgTyrTyrAspLysAsnIleMetThrLysVal 353
Db 1016 TATGACAGCTACCGCGCGCTCCGCTACTACTACACAAAAACATCATGACCAAGTG 1075
QY 354 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 373
Db 1076 CAGGGAGAGCGTACGCTCAAGTTGACTCCACGGGATGCGCCAGCGCTGCACGCC 1135
QY 374 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 393
Db 1136 CACCTTCCTGAGTGCCTCGTACAGTACCCCTCCGACCTGCCATCATGAGGCTCTCTAT 1195
QY 394 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 413
Db 1196 CAGGCCACCCCGAGAGATGATTTGTCTCCACACCTCCGCTCCAGTACA 1255
QY 414 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 433
Db 1256 TCTTCCAGTTCTTTGTTCCCGCAACCATCTGATGATTCACAGCATGGGGCATCTAC 1315
QY 434 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 451

Db 1316 CCGAAGCACTAGCTCCAGCGACCATATGCGCTCAGCTGCGGCGACTACTAC 1369
RESULT 11
PCT-US02-41414-1210
Sequence 1210, Application PC/TUS0241414
GENERAL INFORMATION:
APPLICANT: Morris, David
APPLICANT: Engelhard, Eric
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71249/RMS/DCF
CURRENT APPLICATION NUMBER: PCT/US02/41414
CURRENT FILING DATE: 2002-12-26
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 1613
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1210
LENGTH: 1359
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US02-41414-1210
Alignment Scores:
Pred. No.: 4,85e-178 Length: 1359
Score: 1624.50 Matches: 306
Percent Similarity: 79.43% Conserved: 57
Best Local Similarity: 66.96% Mismatches: 83
Query Match: 66.47% Indels: 11
Gaps: 7
US-09-902-772-2 (1-451) x PCT-US02-41414-1210 (1-1359)
QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 1 ATGAGCGGAGCTATTAAGAGAGCTGTGCTGCTGAGCAGCAGCCAGTCCCTTTGAC 60
QY 21 CysAlaTyrGlySerProHisLeuAlaThrGluMetThrAlaSerSerSerSer 39
Db 61 TCACGCTACGAGCGGAGCGGAGCCATCTCCCAAGCGCAGCATGAGTGGGAGAGCT 120
QY 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
Db 121 GACTACGCGGAGCGCCCAAGATCAACCCCTCCACACAGCAGAGTGAATCAATCAG 180
QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 181 -----CCAGTGAAGGTCAACGTCACAGCGGAGTAT-----GACCACATGATGATCCAGG 231
QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 232 GAGTCTCGGTGAGCTCAGCTCAGCTTACGAATGACGAAAGCTGGTGGAGGAGGAGTCC 291
QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMet 118
Db 292 AACCCCATGAATCAACACCTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 351
QY 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuThrPheSerThrAspHis 138
Db 352 ACACACCAACAGAGAGAGATGATGCTCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAT 411
QY 139 ValArgGlnTrpLeuGluTrpAlaValIleGlyGluTyrGlyLeuProAspValAspIleLeu 158
Db 412 GTGAGCGAATGCTGAGTGGGCGCATTAAGAGATACGCTTGATGAGATGACACATCC 471
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
Db 472 TTTTCCAGAAATGATGCGCAAGAGATGTGTAATAATGAAAGAGAGAGAGAGAGAGAG 531
QY 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgValArg 198
Db 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgValArg 198


```

Db      532 GCCACACACCTCTACAAACAGGAGTGCTGTGTACACACCTTACCTACAGGAA--- 588
Qy      199 GYAlaThrPheIlePheProAsnThrSerValTyProGluAlaThrGlnArgIleThr 218
Db      589 ---AGTTCACCTGCGGCTATTAATACAACTCCACACCGCCACCAATCTCCAGATGAGT 645
Qy      219 ThrArgProAsnLeuProTyArgGlnAlaArgSerAlaThrPheSerHisSerHis 238
Db      646 GTCAAAAGAACCTCTTTATGACTCAGACAGAGAGAGCTTGCGCAATACATGAAAT 705
Qy      239 ProThrGlnSerIleValAlaThrGln---ProSerSerSerThrValProIleTyThrGluAsp 257
Db      706 TCTGCGCTCAACAAAGTCTCCCTTGAGAGGCGCACAAACGATCAGTAAGAAATACAGAG 765
Qy      258 GlnArgProGlnLeuAspProTyArgGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
Db      766 CAACGGCCCGCCAGCCAGATCCGTATCAGATCCTGGGCGCCGACACAGATCCCTGAGCCAC 825
Qy      278 ProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeuLeuLeuSerAspSerSer 297
Db      826 CTTGAAAGCGGCGAGATCCACCTGTGCAATTCCTCTGAGAGTCTCTCCGACAGCGCC 885
Qy      298 AsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyIlePheLeuMetThrAspProAsp 317
Db      886 AACCCAGCTGTATCAGCTGGAGGAGGACCAACGGGAGTTCAAAATGACGAGACCCCGAT 945
Qy      318 GluValAlaArgAlaGlyTrpGlyGlnArgIleSerIleSerIleProAsnMetIleValIle 337
Db      946 GAGGTGGCCAGGCGCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
Qy      338 SerArgAlaLeuArgIleGlyTrpAspIleValIleValIleValIleValIleValIle 357
Db      1006 AGCGGCGCTCCCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065
Qy      358 TyrAlaTyrlsPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
Db      1066 TATGCTTACAAATTTGACTTCACAGGATGCGCAGGCTGCGACGACATCCAGCCAG 1125
Qy      378 SerSerMetTyrlsTyrlsProSerAspLeuProTyrlsMetSerSerTyrlsAlaHisPro 397
Db      1126 TCGCCACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 1185
Qy      398 GlnIleMetAsnPheValAlaProHisProAlaLeuProValThrSerSerSerPhe 417
Db      1186 CAGAGGTGAATCTTGTCCCTCCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 1245
Qy      418 PheAlaAlaProAsnProTyrlsPheAsnSerProThrGlyIleTyrlsProAsn----- 435
Db      1246 TTTGGAGCCGATCAATGACTGAGACCTCCCGCCAGGGGGGAAATCTACCCCAACCCCAAC 1305
Qy      436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrlsTyrls 451
Db      1306 GTCCCCCGCATCTTACACCCAGGCTGCTTACACTTACGACACTTACGACACTTAC 1356

```

RESULT 12

```

PCT-US02-41414-1209
; Sequence 1209, Application PC/TUS0241414
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71249/RMS/DCF
; CURRENT APPLICATION NUMBER: PCT/US02/41414
; CURRENT FILING DATE: 2002-12-26
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1613
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1209
; LENGTH: 2957
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
PCT-US02-41414-1209
Alignment Scores:
Pred. No.: 1,97e-177 Length: 2957
Score: 1624.50 Matches: 306
Percent Similarity: 79.43% Conservative: 57
Best Local Similarity: 66.96% Mismatches: 83
Query Match: 66.47% Indels: 11
DB: 2 Gaps: 7
US-09-902-772-2 (1-451) x PCT-US02-41414-1209 (1-2957)
Qy      1 MetAlaSerThrIleTyrlsGlnAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db      173 ATGAGCGGAGTATTAAGAGGCTCTGCTGTGAGCGAGCACAGTCCCTCTTTGAC 232
Qy      21 CysAlaTyrlsGly---SerProHisLeuAlaAlaThrGlnMetThrAlaSerSerSer 39
Db      233 TCAGCTACGAGAGGCGGAGCCATCTCCCAAGGCCGACATGATGCTCGGGAGTCT 292
Qy      40 GluTyrlsGlnThrSerIleMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
Db      293 GACTACGGGAGAGCCCAAGATCAACCCCTCCACACAGCAGAGAGTGGATCAATCAG 352
Qy      60 ProProAlaArgValThrIleLeuMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db      353 ---CCAGTGAAGGGTCAACGTCACGGGAGTAT---GACCACATGAATGATCCAGG 403
Qy      80 AsnSerProAspAspCysSerValAlaAlaGlyGlyIleTyrlsMetValSerSerSerAspAsn 99
Db      404 GAGTCTCGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 463
Qy      100 ValGlyMetAsnTyrlsGlySerTyrlsMetGluGluHis---IleProProAsnMet 118
Db      464 AACCCATGAATCAACAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 523
Qy      119 ThrThrAsnGlnArgValIleValProAlaAspProThrIleTyrlsSerThrAspHis 138
Db      524 ACCACACAGAGAGAGAGATGATGCTCCCGCCAGACCCACACAGTGTGACACAGAGCAT 583
Qy      139 ValArgGlnTrpLeuGlnIleValAlaValIleGlyIleLeuProAspValAspIleLeu 158
Db      584 GTGAGGCAATGGCTGAGAGTGGCCATTAAGAGATGACAGCTTGATGAGATGACACATCC 643
Qy      159 LeuPheGlnAsnIleAspGlyIleGluLeuGluCysIleMetThrIleAspAspPheGlnArg 178
Db      644 TTTTCCAGAAATGATGAGCAAGACTGTGTAATGAACAAGAGAGAGACTTCTCCGC 703
Qy      179 LeuThrProSerTyrlsAsnAlaAspIleLeuLeuSerHisLeuHisTyrlsLeuArgGluArg 198
Db      704 GCCACACCTCTTACACAGGAGTGTGTGACACCTGATTAATGAGTAACTGAGGAA--- 760
Qy      199 GYAlaThrPheIlePheProAsnThrSerValTyProGluAlaThrGlnArgIleThr 218
Db      761 ---AGTTCACCTGCGGCTATTAATACAACTCCACACCGCCACCAATCTCCAGATGAGT 817
Qy      219 ThrArgProAsnLeuProTyArgGlnAlaArgSerAlaThrPheSerHisSerHis 238
Db      818 GTCAAAAGAACCTCTTTATGACTCAGACAGAGAGAGAGCTTGCGCAATACATGAAAT 877
Qy      239 ProThrGlnSerIleValAlaThrGln---ProSerSerSerThrValProIleTyThrGluAsp 257
Db      878 TCTGCGCTCAACAAAGTCTCCCTTGAGAGGCGCACAAACGATCAGTAAGAAATACAGAG 937
Qy      258 GlnArgProGlnLeuAspProTyArgGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
Db      938 CAACGGCCCGCCAGCCAGATCCGTATCAGATCCTGGGCGCCGACACAGATCCCTGAGCCAC 997
Qy      278 ProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeuLeuSerAspSerSer 297
Db      998 CTGGAAGCGGCGAGATCCAGTGTGCAATTCCTCTGAGAGTGTCTCCGACAGAGGCC 1057

```

QY 298 AsnSerAnCysIleThrTrpGluGlyThrAsnGlyLysIleuPheLysMetThrAspProAsp 317
 Db 1058 AACGGCAGCTGTATACCTGGAGGGGAGCAACGGGAGGTCAAAATGAGGAGCCCGAT 1117
 QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetLanTrpAspLysLeu 337
 Db 1118 GAGGTGGCCAGGGCGGTGGGGGAGGAGGAAAGCAAGCCAAATGAAATTCAGACAACTG 1177
 QY 338 SerArgAlaLeuArgTrpGlyTrpAspLysAsnIleMetThrLysValHisGlyLysArg 357
 Db 1178 AGCGGGCCCTCCCTTTTACTATGATGATTAATAACATTTATGACCAAGTGCAGCGCAAGA 1237
 QY 358 TyrAlaTrpLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
 Db 1238 TATGCTTACAAATTTGACTTCCACGGCATTTGCCAGCTGTGCAGCCATCCAGCCAGAG 1297
 QY 378 SerSerMetLysTrpLysTrpSerAspLysProGlyMetSerLysTrpHisAlaHisPro 397
 Db 1298 TCGTCCATGTACAGATCCCTTCTGACATCTCTTACATGCTTCTTCCATGCCACGAC 1357
 QY 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerPhe 417
 Db 1358 CAGAGGTGAACCTTTGTCCTCCCATCCATTCCTCCATGCTGTCACTTCTCCAGCTTC 1417
 QY 418 PheAlaAlaProAsnProTrpTrpAsnSerProThrGlyLysIleTrpProAsn 435
 Db 1418 TTGGAGCCCGCATCAATCTGACCTCCCGCCAGGGGGAATCTACCCCAACCCCAAC 1477
 QY 436 ---ThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTrpTrp 451
 Db 1478 GTCCCGCCCATCTTACACCCACCGAGTGCCTTTCACATTTAGGACACTTAC 1528

RESULT 13
 US-10-342-887-710
 : Sequence 710, Application US/10342887
 : GENERAL INFORMATION:
 : APPLICANT: Dai, Hongyue
 : APPLICANT: He, Yudong
 : APPLICANT: Linsley, Peter S.
 : APPLICANT: Mao, Mao
 : APPLICANT: Roberts, Christopher J.
 : APPLICANT: Van 't Veer, Laura Johanna
 : APPLICANT: Van de Vijver, Marc J.
 : APPLICANT: Bernards, Rene
 : TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 : FILE REFERENCE: 9301-188-999
 : CURRENT APPLICATION NUMBER: US/10/342,887
 : PRIOR FILING DATE: 2003-01-15
 : PRIOR APPLICATION NUMBER: 60/298,918
 : PRIOR FILING DATE: 2001-06-18
 : PRIOR APPLICATION NUMBER: 60/380,710
 : PRIOR FILING DATE: 2002-05-14
 : PRIOR APPLICATION NUMBER: 10/172,118
 : PRIOR FILING DATE: 2002-06-14
 : NUMBER OF SEQ ID NOS: 2699
 : SEQ ID NO 710
 : LENGTH: 2957
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-10-342-887-710

Alignment Scores:
 Pred. No.: 1.97e-177 Length: 2957
 Score: 1624.50 Matches: 306
 Percent Similarity: 79.43% Conservative: 57
 Best Local Similarity: 66.96% Mismatches: 83
 Query Match: 66.47% Indels: 11
 DB: 14 Gaps: 7

US-09-902-772-2 (1-451) x US-10-342-887-710 (1-2957)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGlnAspGlnSerLeuPheGlu 20
 Db 1118 TATGCTTACAAATTTGACTTCCACGGCATTTGCCAGCTGTGCAGCCATCCGACGAG 1297

Db 173 ATGAGCGGAGCTATTAAAGAGGCTCTGCGGTGTGAGCAGCAGCAACGCTCTTTGAC 232
 QY 21 CysAlaTrpGly---SerProHisLeuAlaLysThrGlnMetThrAlaSerSerSer 39
 Db 233 TCAGCTTACGGAGGGGAGGCCATCTCTCCCAAGGCCACATGACTGCTTGGGAGTCT 292
 QY 40 GlnTrpGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
 Db 293 GACTACAGGGAGCCGCCAAGATCAACCCCTCCACACAGCAGAGGAGTGAATCAATCAG 352
 QY 60 ProProAlaArgValThrIleLysMetGlyCysAsnProAsnGlnValAsnGlySerArg 79
 Db 353 ---CCAGTGAAGGTCACAGTCAAGCGGAGCTAT---GACCAATGAATGATCCAGG 403
 QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
 Db 404 GAGTCTCGGTGGATCTCAGCTGATAGCAATTCAGCAAGCTGGTGGGAGGAGGATCC 463
 QY 100 ValGlyMetAsnTrpGlySerTrpMetGlnLysHis---IleProProAsnMet 118
 Db 464 AACCCCATGAACTACAAACACTATATGAGAGAGAAATGGCCCCCTCTCCCAACATG 523
 QY 119 ThrThrAsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138
 Db 524 ACCACACAGAGAGAGAGATGATCTCCCGCAGACCCACACTGTGTGACACAGAGCAT 583
 QY 139 ValArgGlnTrpLeuGlnTrpAlaValAlaLysGlnTrpGlyLeuProAspValAspIleLeu 158
 Db 584 GTGAGCGAATAGGCTGAGTGGGCGCATTAAGAGATACAGCTGATGAGATGCACATCC 643
 QY 159 LeuPheGlnAsnIleAspGlyLysGlnLysCysLysMetThrLysAspAspPheGlnArg 178
 Db 644 TTTTCCAGAACAGATGATGCAAGAACTGTGTAATGAAACAAAGGAGGACTTCTCCGCG 703
 QY 179 LeuThrProSerTrpAsnAlaAspIleLeuLeuSerHisLeuHisTrpLysArgGluArg 198
 Db 704 GCCACACCTCTTACAAACAGGAAAGTCTGTGACACCTCACTTACCTCAGAGAA--- 760
 QY 199 GlyAlaThrPheIlePheProAsnThrSerValTrpProGluAlaThrGlnArgIleThr 218
 Db 761 ---AGTTCACTGGTGGCTTAATTAACACTCCACACGACCAACCAATCTCCAGTATGAGT 817
 QY 219 ThrArgProAspLeuProTrpTrpGlnAlaArgArgSerAlaTrpThrSerHisSerHis 238
 Db 818 GTCAAAAGAACCTTCTTATGACTCAGTCAGAAAGAGACTTGGGCAATACATGAAT 877
 QY 239 ProThrGlnSerLysAlaThrGln---ProSerSerSerThrValProLysTrpGluAsp 257
 Db 878 TCTGGCTCAACAAAGTCTCTCCCTGGAGGGGCAAAACATCACTAAGATATACGAG 937
 QY 258 GlnArgProGlnLeuAspProTrpGlnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
 Db 938 CAACGGCCCGCAGCAGAGCCGTATCAATCTGGGCGCCGACCAAGATCCCTTACGCCAAC 997
 QY 278 ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuLeuSerAspSerSer 297
 Db 998 CCTGGAAGCGGAGATCCAGCTGTGCAATCTCTCTGAGAGCTCTCCGACAGGCGCC 1057
 QY 298 AsnSerAnCysIleThrTrpGluGlyThrAsnGlyLysIleuPheLysMetThrAspProAsp 317
 Db 1058 AACGGCAGCTGTATACCTGGAGGGGAGCAACGGGAGGTCAAAATGAGGAGCCCGAT 1117
 QY 318 GluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetLanTrpAspLysLeu 337
 Db 1118 GAGGTGGCCAGGGCGGTGGGGGAGGAGGAAAGCAAGCCAAATGAAATTCAGACAACTG 1177
 QY 338 SerArgAlaLeuArgTrpGlyTrpAspLysAsnIleMetThrLysValHisGlyLysArg 357
 Db 1178 AGCGGGCCCTCCGTATTAATGATGATTAATAACATTTATGACCAAGTGCAGCGCAAGA 1237
 QY 358 TyrAlaTrpLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGlu 377
 Db 1238 TATGCTTACAAATTTGACTTCCACGGCATTTGCCAGCTGTGCAGCCATCCGACGAG 1297


```

1  APPLICANT: Engelhard, Eric
2  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
3  FILE REFERENCE: A-71249/RMS/DCF
4  CURRENT APPLICATION NUMBER: PCT/US02/41414
5  CURRENT FILING DATE: 2002-12-26
6  PRIOR APPLICATION NUMBER: US 09/747,377
7  PRIOR FILING DATE: 2000-12-22
8  PRIOR APPLICATION NUMBER: US 09/798,586
9  PRIOR FILING DATE: 2001-03-02
10 NUMBER OF SEQ ID NOS: 1613
11 SOFTWARE: PatentIn version 3.1
12 SEQ ID NO 1206
13
14 LENGTH: 1729
15
16 TYPE: DNA
17
18 ORGANISM: Mus musculus
19
20 DCT-US02-41414-1206

```

3853.895 Million cell updates/sec

27: em_gss_ro

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	1118.5	42.8	827	9	AJ456498	AJ456498 AJ456498
2	1040	45.6	1045	3	BM456833	BM456833 AGENCOURT
3	1016.5	41.6	880	14	BQ233264	BQ233264 AGENCOURT
4	1002.5	41.0	1089	14	BQ212127	BQ212127 AGENCOURT
5	999	40.9	715	12	BG386025	BG386025 602412867
6	985	40.3	949	13	BI558012	BI558012 603240720
8	984	40.3	642	13	BI713036	BI713036 1e01d08.y
9	980	40.3	568	13	BB660034	BB660034 BB660034
9	980	40.1	592	13	BM088063	BM088063 501436.MA
10	975.5	39.9	915	12	BG390291	BG390291 602416255
11	971.5	39.8	701	9	AU136709	AU136709 AU136709
12	970	39.7	629	10	AM948986	AM948986 QVA -FT000
13	966	39.5	1074	14	BQ593977	BQ593977 AGENCOURT
14	948	38.8	675	10	BB639043	BB639043 BB639043
15	947.5	38.8	634	13	BM489636	BM489636 p9m2n..PKO
16	935.5	38.3	928	12	BG259459	BG259459 602378556
17	932	38.1	629	10	BB577887	BB577887 BB577887
18	840.5	34.9	671	13	BI585830	BI585830 603240526
19	828.5	33.9	666	12	BG390572	BG390572 602416344
20	822	33.6	553	12	BG256864	BG256864 6023771114
21	817.5	33.4	1015	9	AL554108	AL554108 AL554108
22	816.5	33.4	524	9	AU794504	AU794504 fc45dl1.y
23	795.5	32.5	645	10	BB656030	BB656030 BB656030
24	793	32.4	522	12	BF566705	BF566705 UT-R-B2OP
25	776.3	31.8	740	9	AJ456506	AJ456506 AJ456506
26	769	31.5	516	12	BG691566	BG691566 340784.BA
27	735	30.1	651	10	BB442803	BB442803 BB442803
28	722.5	29.6	619	10	AM421310	AM421310 f189g06.y
29	703	28.8	674	10	BB238163	BB238163 BB238163
30	701	28.7	420	13	BI181694	BI181694 UNC -P-PN-
31	695	28.4	520	10	BE233460	BE233460 139636.MA
32	668	27.3	1245	14	BQ212133	BQ212133 AGENCOURT
33	649	26.6	1068	14	BQ062565	BQ062565 AGENCOURT
34	642.5	26.3	812	12	BE831331	BE831331 QVA -MT004
35	638	26.1	572	14	BQ232274	BQ232274 AGENCOURT
36	620	25.4	894	14	BQ439561	BQ439561 AGENCOURT
37	616.5	25.2	756	9	AJ443352	AJ443352 AJ443352
38	616.5	25.2	815	9	AJ446143	AJ446143 AJ446143
39	614	25.1	874	14	BQ215477	BQ215477 AGENCOURT
40	611	25.0	465	14	R87572	R87572 ym89g06..x1
41	610.5	24.9	947	10	BQ642921	BQ642921 AGENCOURT
42	607.5	24.9	787	10	BE559913	BE559913 601345462
43	596.5	24.4	587	9	AL598619	AL598619 DKF2p313I
44	595.5	24.0	551	9	AA467723	AA467723 ve02h06..x
45	585.5	24.0	788	9	AJ456869	AJ456869 AJ456869

JOURNAL Unpublished (2002)
 COMMENT Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
 FEATURES Location/Qualifiers
 source
 1. 827
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="8m13r2"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 /note="CB inbred strain"
 BASE COUNT 233 a 236 c 177 g 181 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,21e-103 Length: 827
 Score: 1118.50 Matches: 211
 Percent Similarity: 78.31% Conservative: 20
 Best Local Similarity: 71.53% Mismatches: 43
 Query Match: 45.77% Indels: 21
 DB: 9 Gaps: 3
 US-09-902-772-2 (1-451) x AU456498 (1-827)
 Oy 138 HisValaArggIntPleuGluTPalaValLysGluTyrglyLeuProAspValaSpile 157
 |||||
 Db 5 CACGTGCGCCAGTGGCTGGATGGCCATTAAGGAGTATGATTAAGATGACACCC 64
 158 LeuLeuPheGlnAsnIleAspGlyLysGluLeuGlySylsMetThrLysAspPheGln 177
 |||||
 Db 65 ACCCTCTCCAGAAATGATGATGCAAGAGCTCTGCAAAAGCAAGACAGATGACTCTC 124
 Oy 178 ArgLeuThrProSerTyrglnAlaAspIleLeuLeuSerHisLeuHisTyrglyLeuArglu 197
 |||||
 Db 125 CGAACCAACCTCTCTCAACACACAGAACTGTGTCTCACTCAGTTACCTCAGGGGA 184
 Oy 198 ArgGlyAlaThrPheIlePheProAsnThrSerValTyrglnAlaThrGlnArgIle 217
 |||||
 Db 185 AGTAGCTACAGTGGCTGCTAC---AATCTCCATCCACACAGAGCTCTCTCAGCTT 241
 Oy 218 ThrThrArgProAspLeuProTyrglnAlaAlaArgSerAlaThrThrSerHisSer 237
 |||||
 Db 242 GCCACCAAGAGGTCCCTGTT----- 265
 Oy 238 HisProthrGlnSerLysAlaThrGlnProSerSerThrValProLysThrGluAsp 257
 |||||
 Db 266 -----GCAGGACACAAACCTGACACAGACA-----ACAGAACAG 301
 Oy 258 GlnArgProGlnLeuAspProTyrglnIleLeuGlyProThrSerSerArgLeuAlaAsn 277
 |||||
 Db 302 CAAGGGCTCAGCAGATCCCTATCAAAATCTGGGGCCCACTAGTAGTGGCTTGGCCAT 361
 Oy 278 ProGlySerGlyGlnIleGlnLeuThrPdnPheLeuLeuGlnLeuLeuSerAspSerSer 297
 |||||
 Db 362 CCTGGAGTGGGAGATACAACTATGGAGTTCCTCCGAGTTCGCTGGACAGTTCC 421
 Oy 298 AsnSerAsnGlyIleThrTPGlnGlyLysAsnGlyLysIleLysMetThrAspProAsp 317
 |||||
 Db 422 AACCCAGCTGTATCATATGGAGGAGCCAAATGGGGAATTCAGATGACAGCCAGAT 481
 Oy 318 GlnValAlaArgArgTPrGlyGluArgLysSerLysProAsnMetAsnTyrglyAspLysLeu 337
 |||||
 Db 482 GAATGGACGCGCCCTGGGGAGAACGCAAAAGCCCAACATGAATATGACAGAGCTG 541
 Oy 338 SerArgAlaLeuArgTyrglyTyrglyAspLysAsnIleMetThrLysValHisGlyLysArg 357
 |||||
 Db 542 AGCGGAGCCCTTGATCTATGATGATAAGACATTTATGACAAAGTSCATGGCAAGG 601
 Oy 358 TyrglyAlaTyrglyPheAspPheHisGlyIleAlaGlnAlaLeuGlnProHisProGln 377

Db 602 TATGCTCAAAATTGACTTCATGAGCATTCGCCAGGCTCTCCACCTCATCCCACTAA 661
 Oy 378 SerSerMetTyrglyTyrglyProSerAspLeuProTyrglyMetSerSerTyrglyHisAlaHisPro 397
 |||||
 Db 662 TCATCATATGATCAAGATCATCATCATCATCTCTCTCTCATCATCATCATCATCATCAT 721
 Oy 398 GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerPhe 417
 |||||
 Db 722 CAGAAAGTGAAGCTTGTATCCACCCACACACCTCTCTATGAGCGTCACATCATCATCAT 781
 Oy 418 PheAlaAlaProAsnProTyrglyThrAsnSerProThrGlyGlyIle 432
 |||||
 Db 782 TTCGAGCAGGCTCAGCTTATGAGCCTCTGCTGGAGACATT 826
 RESULT 2
 BM456833
 LOCUS
 DEFINITION
 5', mRNA sequence.
 ACCESSION
 BM456833
 VERSION
 BM456833.1 GI:18505873
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: L14M12347 row: f column: 10
 High quality sequence stop: 669.
 FEATURES
 source
 1. 1045
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5583753"
 /clone_lib="NIH-MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."
 BASE COUNT 300 a 290 c 230 g 225 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.13e-95 Length: 1045
 Score: 1040.00 Matches: 214
 Percent Similarity: 75.24% Conservative: 17
 Best Local Similarity: 69.71% Mismatches: 33
 Query Match: 42.55% Indels: 44
 DB: 13 Gaps: 8
 US-09-902-772-2 (1-451) x BM456833 (1-1045)
 Oy 7 GlnAlaLeuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrglySerPro 26
 |||||
 Db 190 GAACCTTATACAGTTGAGAGAGACACAGCTGTTTATGAGTGTGCTCAGCAACGCCA 249
 Oy 27 HisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyGlnThrSerLys 46

```

Db 250 CACCTGGCTAGACAGAGATACCGCGCTCTCCACGACCTATGACACACTTCCAG 309
Oy 47 MetSerProAlaValProGlnGlnAspTyrLeuSerGlnProProAlaValThrIle 66
Db 310 ATGAGCCACACCGCTCCACAGAGATTGGCTGTCTCAACCCACGAGGTCACACATC 369
Oy 67 LysMetGluCysAsnProAsnGlnValAsnGlySerAspAsnSerProAspCysSer 86
Db 370 AAAATGGATTAACCCACGAGTGAATGCTCAAGAACTCTCTGATGAATGCACT 429
Oy 87 ValAlaLysGlyGlyLysMetValSerSerSerSerAspAsnValGlyMetAsnTyrGlySer 106
Db 430 GTGGCCAAAGCGGGAGAGATGTGGCGACCCACACACCGTTGGATGAACCTACGCGCAGC 489
Oy 107 TyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluValArgValIle 126
Db 490 TACATGGAGGAGAGACATGACACCCCAACATGACACGAGACGAGCGCATATC 549
Oy 127 ValProAlaAspProThrLeuTrpSerThrAspHisValArgGlnTrpLeuGluTyrPala 146
Db 550 GTGGCAGCAGATCTTACGCTATGAGATACACCATGTGGCGCAGTGGCTGGAGTGGCGC 609
Oy 147 ValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166
Db 610 GTGAAGAGATATGGCTCCAGACGTCAACATCTGTATTATCCAGAAACATGATGGAGAG 669
Oy 167 GluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186
Db 670 GAATGTGACAGATGACCCAGACGACCTCCAGAGGCTCACCCACCTACACGCGCAC 729
Oy 187 IleLeuLeuSerHisLeuHisTyrLeuArgGluArgGlyAlaThrPheIlePheProAsn 206
Db 730 ATCTTCTCTACATCTTCTACTACTCAGAGAGACTCTCTTCCACATTGACTCCAGAT 789
Oy 207 ThrSerValTyrProGlnAlaThrGlnArgIleThrThrArgPro----- 221
Db 790 -----GATGTGTGATAAAGC-CTTACAAACCTCCCGGTATATGATCATGCT 833
Oy 222 -----AspLeuProTyrGluGlnAlaAlaArgArgSerAlaTrpHisSerHis 238
Db 834 AGAAACACCGATTTACATATGAGACCCCGGAGATCAGCTGAGCGGTCAACGCTCA 893
Oy 239 ProThrGlnSerLysAlaThrGlnProSerSerSerThrVal-----Pro 253
Db 894 CCC-----CAAGCCCAAGTGAAGAGCTTCAACATCTTCCCTTCAAGGCGCC 944
Oy 254 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 273
Db 945 AAAACTGAA-----AGACCCAGC-----GGCCCTCAGGT--- 974
Oy 274 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrProIleLeuGluLeuLeu 293
Db 975 -----TAAATTCCT-----TTAATCAACATCTT 998
Oy 294 SerAspSerSerAsnSerAsn 300
Db 999 TGGGACACCCACATATTAAC 1019

```

```

RESULT 3
BQ233264 880 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT 7565836 NIH_MGC_92 Homo sapiens cDNA IMAGE:6041526
DEFINITION 5', mRNA sequence.
ACCESSION BQ233264
VERSION BQ233264.1 GI:20414664
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Plimates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 880)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/

```

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMU at:
http://image.lnl.gov
Plate: L14M13279 row: h column: 07
High quality sequence stop: 608.
Location/Qualifiers
1. 880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6041526"
/clone_id="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 250 a 231 c 207 g 189 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 5,98e-93 Length: 880
Score: 1016.50 Matches: 195
Percent Similarity: 84.94% Conservative: 8
Best Local Similarity: 81.59% Mismatches: 7
Query Match: 41.59% Indels: 29
DB: 14 Gaps: 2
US-09-902-772-2 (1-451) x BQ233264 (1-880)
Oy 1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 74 ATGGCCAGCACTATTAAAGGAGCTTATCAGTTGTGAGTGAAGACAGTGGCTGTGGAG 133
Oy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGlu 40
Db 134 TGTGCTACGGAAGCCACACCTGTGTAAGACAGATGACCGGCTCTCTCCAGCGAC 193
Oy 41 TyrGlnLysThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
Db 194 TATGACAGACTTCCAGATGAGCCACGCGCTCCAGCAGATGCGTGTCAACCC 253
Oy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerAsn 80
Db 254 CCAGCCAGGGTCCACATCAAAATGGAATGTAAACCTTGCAGGAGGATGCTCAAGGAC 313
Oy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnVal 100
Db 314 TCTCTGATGATGACATGTGTGGCCAAAGGGGAGATGTGGCGACCCACAGACCGCTT 373
Oy 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
Db 374 GGGATGAACATAGCGCAGCTCATGTGAGAGACACATGACACCCCAACATATACACAG 433
Oy 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
Db 434 AACGAGCGCAGAGTATCTGCCAGCAGATCTTACCTATGAGATACAGACCATGTGCGG 493
Oy 141 GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
Db 494 CAGTGGCTGAGTGGGGGGTGAAGATATGGCTTCCACACGTCAACATCTTGTATATC 553
Oy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180

```


Db 554 CAGAACATCGATGGGAGGAACCTGTGCAGATGACCAAGACGACTTCCAGAGGCTCACC 613
 QY 181 ProserThyransAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgLysGluArg----- 198
 Db 614 CCCAGCTACAGACGGCGACATCTCTCTCAGATCTCCACATCTCAGAGAGACTCTCTT 673
 QY 198 ----- 198
 Db 674 CCACATTTGACTTCAGATGATGNTGATAAGCCCTTACAAAACCTCCACGCGGNTAATGC 733
 QY 199 -----GlyAlaThrPheIlePheProAsn---ThrSerValTyr 210
 Db 734 ATGGCTAGAAACACAGAGGGGGGTGCGAGCTTTTATTATTTCCCAATTAATCTCAGTATAT 790
 RESULT 4
 B0212127 1089 bp mRNA linear EST 02-MAY-2002
 LOCUS AGENCOURT_7571287 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6052136
 DEFINITION B0212127
 ACCESSION B0212127
 VERSION B0212127.1 GI:20392058
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1089)
 NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@ps-femail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM13307 row: b column: 09
 High quality sequence stop: 562.
 Location/Qualifiers
 1..1089
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6052136"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 292 a 314 c 257 g 225 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.2e-91 Length: 1089
 Score: 1002.50 Matches: 207
 Percent Similarity: 70.31% Conservative: 18
 Best Local Similarity: 64.69% Mismatches: 45
 Query Match: 41.02% Indels: 50
 DB: 14 Gaps: 5
 US-09-902-772-2 (1-451) x B0212127 (1-1089)
 QY 1 MetaLaserThrllelYsgLualAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 76 ATGGCCAGCACTATTAAAGAGCCTTATCAGTTGTGAGTGAGCAGCAGTCCTTTGAG 135
 QY 21 CysAlaIarYrGlySerProHisLeuAlaIalYsThrGluMetThrAlaSerSerSerGlu 40

Db 136 TGTCCTACAGGAGACCCACACACTGGCTTAAGACAGAGATGACCCGCTCTCTCCAGGAC 195
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60
 Db 196 TATGACAGACTTCCAAAGATGAGCCACAGCGCTCCCTCANAGATTTGGCTGTCAACCC 255
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 256 CCAGCCAGGGGTCACCATCAAAATGAAATGTAACCTAGCCAGGATGATGCTCAAGGAC 315
 QY 81 SerProAspAspCysSerSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnVal 100
 Db 316 TCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProPheProAsnMetThrCys 120
 Db 376 GGGATGAACCTACGCGACCTACATGAGAGAGAGACATGCCACCCCAACATACACACG 435
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
 Db 436 AACGAGCGCAGAGATTATCTGTCAGAGATCTTACGCTATGAGTACAGACATGTCGG 495
 QY 141 GlnThrLeuGluThrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 Db 496 CAGTGGCTGGAGTGGGGGGAAGAAATATGCGCTTCAGACGTCACATCTTGTATTC 555
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 556 CAGAACATCCATGGGAAGAACTGTGCAAGATGACCAAGCAGACTTCCAGAGGCTCACC 615
 QY 181 -ProserThyransAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgLysGluArg----- 198
 Db 616 CCCAGCTACATGGCCGACATCTCTCTCAGATCTCAGTCTCAGAGAGACTCTCT 675
 QY 198 ----- 198
 Db 676 TCACATTTGACTACATGATGCTGTGATAAGCTTTCAAAACCTCCACAGGCTAATGC 735
 QY 199 -----GlyAlaThrPheIlePheProAsnThr-----SerValTyr 210
 Db 736 ATGGCTAAACCCAGGGGGGCGGCGAGCTTTTATTTCCCAATTCAGTCCAGATATTCCT 795
 QY 210 rProGluAlaThrGlnArgIleThrThrArg-ProAspLeuProTyrGluGlnAlaArg 230
 Db 796 GGAAGCTACGCCCAAAATTTACACCTAGAGGCCAAATTTACCTTATGAGCCCCCCC 855
 QY 230 rg-----SerAlaThrThr-SerHisSerHisProThrGlnSerLysAlaThrGlnPro 247
 Db 856 GGAAGATCACGCCCTTGACCGGTCATGGCCACCC----- 891
 QY 248 SerSerSerThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIle 267
 Db 892 -----CCGAGCGCCGAGGTCGAAAGAGCGTGGCCCCCCCAATTC----- 933
 QY 268 LeuGlyProThrSerSerArgLysLeuAlaAsnProGlySerGlyGlnIleGlnLeu 285
 Db 934 ---TCCCGGGGGGCCCAAACTGAAAGAACACAGCGGTGCCCTGTGTTAAATC 984
 RESULT 5
 Bg388025 715 bp mRNA linear EST 12-MAR-2001
 LOCUS 602412867F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:45123 5'
 DEFINITION mRNA sequence.
 ACCESSION Bg388025
 VERSION Bg388025.1 GI:13281471
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 715)
 NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

**JOURNAL
COMMENT**

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0420 row: f column: 18
High quality sequence stop: 693.

BASE COUNT	206 a	188 c	172 g	149 t
ORIGIN				

Alignment Scores:	
Pred. No.:	2,64e-91
Score:	999.00
Percent Similarity:	97.468
Best Local Similarity:	94.428
Query Match:	40.888
DB:	12
Length:	715
Matches:	166
Conservative:	6
Mismatches:	5
Indels:	0
Gaps:	0

[illegible]

QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlu 197
 |||||
 Db 634 CCCAGCTACACGCCGACATCCTCTCTCACATCTCCAAATACCTCAGAG 684

LOCUS	BI558012	949 bp	mRNA	linear	EST 05-SEP-2001
DEFINITION	603240720F1 NCI_CGAP_Mam4 Mus musculus cDNA IMAGE:5293492 5' , , ,				
ACCESSION	mRNA sequence.				
VERSION	BI558012				
KEYWORDS	BI558012.1 GI:15445326				
SOURCE	EST.				
ORGANISM	house mouse.				
	Mus musculus				

Plate: LAM11742 row: h column: 05
High quality sequence stop: 876.

```

/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:5293492"
/clone_id="NCI CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; site.1: salI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

```

Alignment Scores:			
Pred. No.:	1.08e-89	Length:	949
Score:	985.00	Matches:	190
Percent Similarity:	79.71%	Conservative:	30
Best Local Similarity:	68.84%	Mismatches:	47
Query Match:	40.30%	Indels:	9
DB:	13	Gaps:	3

US-09-902-772-2 (1-451) x B1558012 (1-949)	
QY	183 YTRASNALASPILLeLeuEuserSertSLeuhISYrLeAurGguLuarGJyAlaThrPhe 202
	: : : : : : : : : :
Db	2 TACACACAGGAAGTGTCTGTTCGCACTCACTTACCTCAGGAA-----AGTTCACGTG 55
QY	203 lAepheProAsnTrSertValYrPProGuaLathrGuaAglIerThrHarPProAsP 222
	::: ::: : : : : : ::: : : : : :
Db	56 CTGGCCTTAATACACAACACTCCCATACAGACCAGTCCCTCAGCAGTGAATGCAAGAGAC 115
QY	223 LeuProTYglGluGlnAlaArgSerAlaTrpThrSerHisSerHisPProThGlnSer 242
	: : : : : : : : : :
Db	116 CCTCTCTTAAGACTCTGTCTCAGGAGAGACAGGAAGCAATAATATAGCACTTGCGCTTACC 175

QY	243	lysa1a	-----ThrglnProserSerThrVal1ProLyThrGlnAspGlnAprrogin	261
Db	176	AAAAGTCTCTCTCCCTTGGAGAGATACAGACCATGGGCAAGAACACTGAGCAGCGGCCAG	235	
QY	262	leuaaspprotyrGln1leleuglyProThrSerSerArgleuAlaAsnProglySergly	281	
Db	236	CCAGATCCCTTATCAATCATCTGGGGCCCAACACAGACCCGCTAGCAAACTTGGAGTGG	295	
QY	282	Gln1leGlnleuThrGlnPheleuenu1leuleuSerAspSerSerAsnSerAsn	301	
Db	296	CAGATCCAGCTGTGGCAGTTCCTCTGGAAACTACGTCTCCGACAGGCCAGCGCAGT	355	
QY	302	1leThrTrpGlnGly1ThrasnGly1uNhelysmetThrAspProAspGlyValAlaArg	321	
Db	356	ATCACCCTGGAGGGAGCCACACGGGAGTTCAAAATGACGAGACCTGATGAGTGGCCAGG	415	
QY	322	ArgTrpGlyGlnArglysSerlysProAsnMetAsnTrpAspLysleuSerArgAlaLeu	341	
Db	416	CGCTGGGGAGAGCGGAGAACGCCCAACATGATTTATGACAACTGACCGGGCCCTC	475	
QY	342	ArgTrpTrpTrpArgAspLysAsn1leuethrLysVal1hisGlyLysArgTrpAlaTrpLys	361	
Db	476	CGATCTCTATGACAAAAACATTATGACCAAAAGTCATGGCAAAAGTATGCTACACAG	535	
QY	362	PheaspphehisGly1leAlaGlnAlaLeuGlnProhisAspProGlnLysSermetTrp	381	
Db	536	TTTGACTTCATGGATGGATGGCCAGGCCCTGCAGCCACATCCACAGACATCCATGTAC	595	
QY	382	LysTrpProSerAspLeuProTrpMetSerSerTrpHisAlaHisProGlnLysMetAsn	401	
Db	596	AAGTATCCCTGTGATATCTCTACATGCGCTTCCTACATGCCCATCAACAGAGGTGAAC	655	
QY	402	PheValAlaProhisAspProAlaLeuProValThrSerSerSerPhePheAlaAlaApr	421	
Db	656	TTTGTCGCCGTCACCCCATCTCCATGACCTGTGCACCTCTCCAGCTCTTTGGAGCAGC	715	
QY	421	oAsnProTrpTrpAsnSerProThrGlyGly1leTrpProAsn-----ThrArgL	438	
Db	716	ATCACAAATAGATTGGACCTGCCCACTGCTGGGATCTATCAAAOCCCACTGCCCGCC	775	
QY	438	eupProAlaAlaHisMetProSerHisleuGlyThrTrpTrp 451		
Db	776	ATCCTAACACCCAGCTGCTTTTCACTTAGGACAGCTACTAC 817		
RESULT 7				
B1713036		B1713036	642 bp	mRNA linear EST 11-MAR-2002
LOCUS		leoid08.y1	HR85	islet Homo sapiens cDNA clone IMAGE:5085758 5'
DEFINITION				similar to SW_ERG_HUMAN P11308 TRANSCRIPTIONAL REGULATOR ERG ;
ACCESSION		B1713036		mRNA sequence.
VERSION		B1713036.1	GI:15688731	EST.
KEYWORDS				
SOURCE		human.		
ORGANISM		Homo sapiens		
REFERENCE		Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
		1 (bases 1 to 642)		
		Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,		
		Lemishka,I., Scaerza,M., Brestelli,J., Gradwohl,G., Clifton,S.,		
		Hillier,T., Marra,M., Page,D., Wylie,T., Martin,J., Blisstein,A.,		
		Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas		
		,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.		
		, Jackson,Y. and Bowers,Y.		
		Endocrine Pancreas Consortium		
		Unpublished (2000)		
TITLE		Other_ESTs: leoid08.x1		
JOURNAL		Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue		
COMMENT		Endocrine Pancreas Consortium		
		Harvard University, Howard Hughes Medical Institute		
		Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,		
		MA 02138		

FEATURES		SOURCE	
		Location/Qualifiers	
1. 642		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5085758"	
		/clone.lib="HR85.islet"	
		/tissue.type="Purified pancreatic islet"	
		/lab_host="DH10B"	
		/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1	
		NotI; Site:2: XhoI; cDNA made by oligo-dT priming.	
		Size-selected on agarose gel. Average insert size ~1kb. 5	
		XhoI site was destroyed after directional cloning.	
		Amplified once. Contact information: Hiroshi Inoue, MD,	
		Metabolism Div. (Alan Permut lab), Washington University	
		School of Medicine, Box 8127, 660 South Euclid Ave., St.	
		Louis, MO 63110, E-mail: hinoue@igate.wustl.edu, Tel:	
		314-362-1916, Fax: 314-747-2692."	
BASE COUNT	178 a 171 c 163 g 129 t	1 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	7.57e-90	Length:	642
Score:	984.00	Matches:	183
Percent Similarity:	97.46%	Conservative:	9
Best Local Similarity:	92.89%	Mismatches:	5
Query Match:	40.26%	Indels:	0
DB:	13	Gaps:	0
US-09-902-772-2 (1-451) x BT173036 (1-642)			
QY	1	MetAlaSerThrIleIysGluAlaLeuSerValIserGluAspGlnSerLeuPheGlu	20
DB	52	ATGGCCGACACTATTAAGAGAGCCCTATCAGTTGGAGGAGGACCACTGCTTTGAG	111
QY	21	CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu	40
DB	112	TGTGCTACGGAAGCGCACACCTGCTTAAGACAGAGATGACCGGCTCTCCACGAGAC	171
QY	41	TyrGlyGlnThrSerIysMetSerProAlaGlyLProGlnGlnAspThrLeuSerGlnPro	60
DB	172	TATGAGACAGCTCCAAAGATGAGCCACGCGCTCCACGAGAGTGGCTTCTCAACC	231
QY	61	ProAlaArgValThrIleIysMetGluCysAsnProAsnGlnValaAsnGlySerArgAsn	80
DB	232	CCAGCCAGGGTCCATCAAAATGAGATTAACCTTACCAGGTGAATGGCTCAAGGAC	292
QY	81	SerProAspAspCysSerValAlaIysGlyGlyIysMetValSerSerSerAspAsnVal	100
DB	292	TCTCCGATGATGAATGCGATGTGGCAAAGCGGGAGATGGTGGCAGCCAGACCGTT	351
QY	101	GlyMetAsnTyrGlySerTyrMetGluGluIysHisIleProProAsnMetThrThr	120
DB	352	GGGATGACTACGAGCGACTACATGGAGGAGGAAGCAATGCCACCCCAAAACATGACACG	411
QY	121	AsnGluArgValIleValProAlaAspProThrIleuThrAspHisValArg	140
DB	412	AACGACGAGAGATTATCTGTCACACATCTTACGCTATGGATGATACAGACATGTCCG	471
QY	141	GlnThrLeuGlnThrPalaValaLysGlyIuThrGlyLeuProAspValaAspIleLeuPhe	160
DB	472	CAGTGGCTGAGTGGCGGTGAAGAATATGGCTTCCAGACGTCACATCTTGTATTC	531
QY	161	GlnAsnIleAspGlyIysGluLeuCysIysMetThrIysAspAspPheGlnArgLeuThr	180

Db	532	CAGAAACATGCATGGGGAGCGACCTGTGCAGATGAGCCAGAGACGACTTCCAGAGCTCAC	591
Oy	181	ProSerTyrrAsnAlaaspIleLeuLeuSerHisLeuHisTyrLeuArgLu	197
Db	592	CCGACGTANACGCGCGACATCCTTCTCTCATCATCTCCACATCCACAGAA	642
RESULT 8			
LOCUS	BB660034		
DEFINITION	BB660034	668 bp	mRNA
ACCESSION	BB660034	RIKEN full-length enriched, 13 days embryo	EST 26-OCT-2001
VERSION	BB660034	musculus cDNA D43027L04 5', mRNA sequence.	Lung Mus
KEYWORDS	BB660034		
SOURCE	BB660034.1	GI:16493855	
ORGANISM	EST.		
REFERENCE	house mouse.		
AUTHORS	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 668)		
	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,		
	Himoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,		
	M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,		
	Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,		
	D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,		
	Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,		
	Muramatsu,M. and Hayashizaki,Y.		
	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
TITLE	Unpublished (2001)		
JOURNAL	Contact: Yoshihide Hayashizaki		
COMMENT	Laboratory for Genome Exploration Research Group, RIKEN Genomic		
	Sciences Center(GSC), Yokohama Institute		
	The Institute of Physical and Chemical Research (RIKEN)		
	1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan		
	Tel: 81-45-503-9222		
	Fax: 81-45-503-9216		
	Email: genome-res@gsc.riken.go.jp,		
	URL: http://genome.gsc.riken.go.jp/		
	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh		
	M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new		
	genes. Genome Res. 10 (10), 1617-1630 (2000)		
	wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,		
	Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura		
	S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and		
	Hayashizaki,Y.		
	RIKEN integrated sequence analysis (RISA) system-384-format		
	sequencing pipeline with 384 multicapillary sequencer. Genome Res.		
	10 (11), 1753-1771 (2000)		
	Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara		
	Y. and Hayashizaki,Y.		
	Computer-based methods for the mouse full-length cDNA		
	encyclopedia: real-time sequence clustering for construction of a		
	nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)		
	Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamana,K. and		
	K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and		
	Hayashizaki,Y.		
	Computational Analysis of Full-length Mouse cDNAs Compared with		
	Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)		
	Please visit our web site (http://genome.gsc.riken.go.jp) for		
	further details.		
	e mouse tissues.		
FEATURES	Location/Qualifiers		
source	1..668		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="D43027L04"		
	/clone_lib="RIKEN full-length enriched, 13 days embryo		
	lung"		
	/tissue_type="Lung"		
	/dev_stage="13 days embryo"		
	/lab_host="DH10B"		
	/note="Site.1: SalI; Site.2: BamHI; cDNA library was		

			prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGACGACGGCCGCATCAGATTCTTTTGTTCVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGGAGACGATTCTCGAATTATTAATTATATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."		
BASE COUNT	184 a	161 c	175 g	146 t	2 others
ORIGIN					
Alignment Scores:					
Pred. No.:	8.04e-90	Length:	668		
Score:	984.00	Matches:	183		
Percent Similarity:	96.94%	Conservative:	7		
Best Local Similarity:	93.37%	Mismatches:	6		
Query Match:	40.26%	Indels:	0		
DB:	10	Gaps:	0		
US-09-902-772-2 (1-451) x BB660034 (1-668)					
Oy	1	MetaLasErThrIleTySGLAlAleuSeRValVaISerGIuASpgLSeRLeuPheglu	20		
Dd	80	ATGCCACCACTATTAAAGAGGCTTGTCATTTGTGAGCGAGCACACTCATTTTGAG	139		
Oy	21	CysAlaTryGIySerProHIsLeuAlaLyShrGlUmeThrAlaSeRSerSeRgu	40		
Dd	140	TGTGGCTACGGAGACGCCACACCCTGGCTTAAGACAGAGATGACCGCATCTCTCCAGTGAC	199		
Oy	41	TyrGlYglInThrSeRlysMetSeRProArGYalProGlnGlnAsPTlPLeuSeRglnPro	60		
Dd	200	TATGGCCAGACATCCAAAGATGATGCCAGATCCTCCAGCAGGAGCATGGCTCTCAAGCC	259		
Oy	61	ProHlaArGYalThrIlelySMetGluNySaSnProBaNgInVnAlaSgLYSeRArGsn	80		
Dd	260	CCAGCCAGGTCACCATCAAGATGAGTAGCAACCTACTGCAAGTGAATGGTTCCAGGAAc	319		
Oy	81	SerProASpAPCySerSeRVAlAlaLySGLYlYSmetValSerSeRASPasnVal	100		
Dd	320	TCACCTGATGAGTGCAAGTGTACAAGAAGTGGAGAGTGGCGACGCCGATATCTG	379		
Oy	101	GlYMeChsAnTYrGIySeRTYMeTglULyShSIlePRoPROPaSmEthrThr	120		
Dd	380	GCGATGAGCTACGGCAGCTACATGAGAGAGCATGTGCCCTCCCAAATATCACACA	439		
Oy	121	AsngluArGYalVAlIleValProAlasprOthrIeuTPSerThrASPhISValArg	140		
Dd	440	AATGAGGCCAAGATGATGCTCCGCAATCTCTGtGAGAGCACAGACCAATGTCGA	499		
Oy	141	GlnTrPLEugInUrPalAlaValLySGlUnTYrGIyeUPRoASpValASpLIleUeudApHe	160		
Dd	500	CAGTGGCTGGAGTGGCGGCTGAAGAATATGGCCNCCANTGtGGAGAGCTTACTATTT	559		
Oy	161	GlnASnIIeaSPgLYlYsgLUleucYSLySMethrILySaspASpHeGlnARgLeuthr	180		
Dd	560	CAGATATATCGATGGGAAGAGACTGTGCMAAGATGACAAAGATGACTTCCACGCGCTCAG	619		
Oy	181	ProSeRYrASnAlaSpLIleUeUSeRISHisUHTyrlenuArg	196		
Dd	620	CCGAGCTPAACATGCCGACATTCTCTCAATCTNCACATCACTCAGA	667		
RESULT 9					
LOCUS	BM08063	592 bp	mRNA	linear	EST 19-NOV-2001
DEFINITION	501436 MARC ZBOV Bos taurus CDNA 5', mRNA sequence.				
ACCESSION	BM08063				
VERSION	BM08063.1	GI:16998691			

KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus

REFERENCE
AUTHORS
1 (bases 1 to 592)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Perte,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keel,J.W.

TITLE
Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle

JOURNAL
MEDLINE
Genome Res. 11 (4), 626-630 (2001)

COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAGC
Plate: 139 row: D column: 21
Seq primer: ATTATGCTGACACTATG.

FEATURES
source
1..592
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site.1: NotI; Site.2: SalI; library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

BASE COUNT
ORIGIN
152 a 177 c 159 g 104 t

Alignment Scores:
Pred. No.: 1,71e-89 Length: 592
Score: 980.00 Matches: 181
Percent Similarity: 96.89% Conservative: 6
Best Local Similarity: 93.78% Mismatches: 0
Query Match: 40.10% Indels: 0
DB: 13 Gaps: 0

US-09-902-772-2 (1-451) x BM088063 (1-592)

QY 1 MetAlserThrIlelysguAlalauserValserGlnaspGlnserleupheGln 20
DB 12 ATGGCCAGCAGCTATTAAAGGCGTATCATGCTGTGAGCAGACAGTCACTGTTGAG 71
QY 21 CysAlATyGlyserProHIsleuAlATyThrGlnumethrAlaserSerSerGln 40
DB 72 TGGCGCTACGGCAGCGCCGACCTGCCACAGACGACATGACCGGCTCTCTCGGGGAC 131
QY 41 TyrGlyGlnThrSerIyMetSerProArgValProGlnInaSPTrpleuSerGlnPro 60
DB 132 TATGACAGACATCCAAAGATGAGCCGCGCGTCCCTCAGCAGAGACTGGCTGTCAACC 191
QY 61 ProAlaArgValThrIlelyMetGluCysAsnProAsnGlnValAsnGlyserArgAsn 80
DB 192 CCAGCCAGGCTCACATCAAGATGATGATACCTTAAACCAAGCTGAACGCTCAAGAAC 251
QY 81 SerProAspAspCysSerValAlaIATyGlyGlyMetValSerSerSerAspAsnVal 100
DB 252 TCCCTGATGATGATGACATGCGCAAGAGTGGGAAGATGGTGGGACACCCCGACACCGTC 311

QY 101 GlyMetAsnTyrglySerTyrmecGluGluLysHisIleProProAsnMetThrThr 120
DB 312 GGGATGAACTACAGACACTACATGAGAGAGACATGTCACACCCCAACATGACACACC 371
QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuThrSerThrAspHisValArg 140
DB 372 AACGAGCGCGCGTCATGCTCCTGCGAGATCCACGCTATGAGTACAGACATGTCGG 431
QY 141 GlnThrleuGluThrPalaaValIlysguTyrglyLeuProAspValAspIleleuPhe 160
DB 432 CAGTGGCTGGAGTGGGAGATCAAGATACGGGCTTCGGACGTGACATCTTATATTC 491
QY 161 GlnAsnIleAspGlyIyGluLeuCysLysMetThrIyAspAspPheGlnArgLeuThr 180
DB 492 CAGATATTCACGCGAGGAGGAGCTGTGCATGACCAAGATGACACTTCCAGAGCTCAC 551
QY 181 ProSerTyraAsnAlaAspIleleuSerHisIleuHis 193
DB 552 CCGAGCTACAAAGCTGACATCTTCTGTGCGACCTCCAC 590

RESULT 10
BG390291 915 bp mRNA linear EST 12-MAR-2001
LOCUS
DEFINITION
602416255F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524852 5', mRNA sequence.

ACCESSION
BG390291
VERSION
KEYWORDS
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM10429 row: m column: 13
High quality sequence stop: 820.

FEATURES
source
1..915
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4524852"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT
ORIGIN
246 a 252 c 227 g 190 t

Alignment Scores:
Pred. No.: 9.53e-89 Length: 915
Score: 975.50 Matches: 215
Percent Similarity: 77.36% Conservative: 14
Best Local Similarity: 72.64% Mismatches: 42
Query Match: 39.91% Indels: 28
DB: 12 Gaps: 5

US-09-902-772-2 (1-451) x BG390291 (1-915)

```

QY      1 MetAlaSerThrIleuylSGuAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      59 ATGGCCAGCAGCTATTAAAGAGCCCTTAATGAGTTGTGAGTAGAGACCGAGTGTGTTGAG 118
QY      21 CysAlaTyrGlySerProHISLeuAlaValThrGluMetThrAlaSerSerSerGlu 40
DB      119 TGTGCCCTACGGAAGCCACACCTGGCTAAGACAGAGTAGACCGCGTCTCTCTCCAGCGAC 178
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60
DB      179 TATGACAGACTTCCAGATGAGACCCAGCGTCCCTCAGCAGATGGCTGTCTCAACCC 238
QY      61 ProAlaArgValThrIleuLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      239 CCAAGCAGGGCTCACCATCAAAATGAATGATACCCCTAGCCAGTGAATGGCTCAAGGAC 298
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      299 TCTCCTGATGAATGCACTGTGGCCAAAGGGGGAAGATGGTGGGACCCACAGACCGTT 358
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB      359 GGGATGAACTACGGCAGCTACATGAGAGAGACACATGCCACC-CCAAACATGACCACG 417
QY      121 AsnGluArgArgValIleValProAla-AspProThrLeuThrSerThrAspHisValAla 140
DB      418 AACGAGCGCAGAGTTATCGTCCAGCAAGATCCTACCTATGAGTACACACATGTCG 477
QY      140 GGTInTPLeuGluTTPAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB      478 GCAGTGGCTGGAGTGGCGGTGAAGAATATGGCTTCCAGAGCTCAACATCTTGTATT 537
QY      160 eGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuTh 180
DB      538 CCAGAACATCGATGGAGGAAAGACTGTGCAGATGACCAAGAGACGACTCCAGAGGCTCAC 597
QY      180 IProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgLysArg---G1 199
DB      598 CCCAGCTACAAATGCGCAGCAT-CTTCTCTCAATCTCTCACTACCTCGAAGACTCCCTTC 656
QY      199 ValAlaThrPheIlePhePro-----AsnThrSerValTyrProGlnValAlaThrGlnArg 217
DB      657 TTCACACTTTCAGCTTCAGATGATGTTGATTAAACCTTAACAATCTCCACGGTTATGC 716
QY      217 eThrThrArgProAspLeuProTyrGluGlnAlaAlaArgSerAlaThrThrSerHisSe 237
DB      717 ATGTAGAAACACAGATTTACATATGAGCCCCCAGAGAGATCAAGGCTGAGCGTCCAGG 776
QY      237 rHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGlu-- 256
DB      777 -CACCCACGCCAGTCACAGTGTACATCTCTTCACTGCGCAAAATGAGACAGGCTCT 835
QY      257 -----AspGlnArgProGlnLeuAspProTyr 265
DB      836 CGGAATCTATGATTTCTGGCACAGTTCCGCTGCATCGGGGGCGCAGCTTGATGCC-- 893
QY      265 rGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySer 280
DB      894 -----GGCGCGGAGC-----CCGCGCTCC 914

RESULT 11
LOCUS   AUI36709 701 bp mRNA linear EST 02-AUG-2002
DEFINITION AUI36709 PLACE1 Homo sapiens cDNA clone PLACE1004911 5', mRNA
sequence.
ACCESSION AUI36709
VERSION   AUI36709.1 GI:10997248
KEYWORDS  EST.
SOURCE    human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 701)
AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
           Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
           Isogai,T.
TITLE      HRI human cDNA project
JOURNAL    Unpublished (2000)
COMMENT    Contact: Takao Isogai
           Genomics Laboratory
           Helix Research Institute
           1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
           Tel: 81-438-52-3975
           Fax: 81-438-52-3986
           Email: genomics@hri.co.jp
           HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
           Research Institute; cDNA library construction: Department of
           Virology, Institute of Medical Science, University of Tokyo, and
           Helix Research Institute.
FEATURES   Location/Qualifiers
           source          1..701
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="PLACE1004911"
                        /clone_1bp="PLACE1"
                        /issue_type="placenta"
                        /note="Vector: PME18SF13"
BASE COUNT 199 a 185 c 166 g 145 t 6 others
ORIGIN
Alignment Scores:
Pred. No.: 1,62e-88 length: 701
Score: 971.50 Matches: 185
Percent Similarity: 94.15% Conservative: 8
Best Local Similarity: 90.24% Mismatches: 10
Query Match: 39.75% Indels: 2
DB: 9 Caps: 1
US-09-902-772-2 (1-451) x AUI36709 (1-701)
QY      1 MetAlaSerThrIleuylSGuAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      85 ATGGCCAGCAGCTATTAAAGAGCCCTTAATGAGTTGTGAGTAGAGACCGAGTGTGTTGAG 144
QY      21 CysAlaTyrGlySerProHISLeuAlaValThrGluMetThrAlaSerSerSerGlu 40
DB      145 TGTGCCCTACGGAAGCCACACCTGGCTAAGACAGAGTAGACCGCGTCTCTCCAGCGAC 204
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60
DB      205 TATGACAGACTTCCAGATGAGACCCAGCGTCCCTCAGCAGATGGCTGTCTCAACCC 264
QY      61 ProAlaArgValThrIleuLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB      265 CCAAGCAGGGCTCACCATCAAAATGAATGATACCCCTAGCCAGTGAATGGCTCAAGGAC 324
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB      325 TCTCCTGATGAATGCACTGTGGCCAAAGGGGGAAGATGGTGGGACCCACAGACCGCTT 384
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB      385 GGGATGAACTACGGAAGCTACATGAGAGACAGACATGCCACCCCAACATGACCACG 444
QY      121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValAla 140
DB      445 AACGAGGCGCAGAGTTATCGTCCAGCAGATCTTACGCTATGAGTAGACAGACATGTGCGG 504
QY      141 GGTInTPLeuGluTTPAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB      505 CAGTGGCTGGAGTGGGGGGGGAAGATATGAGCTTCCACAGCTCAACATCTTGTATTTC 564
QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThr-LysAspAspPheGlnArgLeuTh 180
DB      565 CAGAACATCGATGGAGAAAGAACTGTGCAAGATGACCAAGAGACGACTTCAGANGCTCAA 624

```

OY 180 PProSerTyraAlaAspIleuLeuSerHisIleuHisIleuArgGluArgGlyAl 200
 Db 625 CCCAGCTACACCCGACATCTCTCTCAGATCTCAGTACCTC---ANGAGAGACTC 681
 OY 200 aThrPheIlePhe 204
 Db 682 CTCCTTCACATT 694

RESULT 12
 AM948986 629 bp mRNA linear EST 31-MAY-2000
 LOCUS QV4-FR0005-110500-201-F02 FR0005 Homo sapiens CDNA, mRNA sequence.
 ACCESSION AM948986
 VERSION AM948986.1 GI:8126760
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 629)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV4-FR0005-110
 500-201-f02&t3=2000-05-11&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 628.
 Location/Qualifiers
 1..629
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="FR0005"
 /dev_stage="Adult"
 /note="Organ: prostate,tumor; Vector: puc18; Site_1: Smai;
 Site_2: Smai; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 166 a 173 c 171 g 119 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 1.96e-88 Length: 629
 Score: 970.00 Matches: 180
 Percent Similarity: 97.40% Conservative: 7
 Best Local Similarity: 93.75% Mismatches: 5
 Query Match: 39.69% Indels: 0
 DB: 10 Gaps: 0
 US-09-902-772-2 (1-451) x AM948986 (1-629)
 OY 6 LysGluLeuSerValIleSerGluAspGlnSerLeuPheGluGlyAlaTyrGlySer 25
 ::

Db 19 CAGGAGCCCTTATGACTGTGAGTGGAGGACGACGTCGTTGTTGAGTGGCTACGGAAG 78
 OY 26 ProHisLeuAlaTyrGluMetThrAlaSerSerSerSerGlyGlnThrSer 45
 Db 79 CCACACCGCGCTAAGACAGATGACCGCGCTCTCTCCAGCAGCTATGACAGACTTCC 138
 OY 46 LysMetSerProAlaGValProGlnGlnAspTyrPleuSerGlnProProAlaArgValThr 65
 Db 139 AAGATGAGCCGACCGCGCTCCACAGAGATTGGCTGTCTCAACCCGACGAGGTCACC 198
 OY 66 IleLysMetGluCysAsnProAsnGlnValAsnGlySerAlaGlnSerProAspCys 85
 Db 199 ATCAAAATGCAATTAACCTACGACGTAATGCTCAAGAACTCTCTGATGAATGC 258
 OY 86 SerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGly 105
 Db 259 AGTGTGGCCAAAGCGGAGATGCTGGGCGAGCCGACACCGTTGGGATGAACCTACGGC 318
 OY 106 SerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsnGluArgVal 125
 Db 319 AGCTACATGGAGGAGAGACACATGCCCCCAACATGACGACGACGAGAGATT 378
 OY 126 IleValProAlaAspProThrLeuTrpSerThrAspHisValArgGlnTrpLeuGluTrp 145
 Db 379 ATCTGTCAGCAGATCTCTACGCTATGGAGTACACACATGTGCGGACGTGGCTGAGTGG 438
 OY 146 AlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGly 165
 Db 439 GCGGTGAAGAATATGGCGCTTCCAGACGTCGAACCTGTATGCCAGAACATGATGGG 498
 OY 166 LysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAla 185
 Db 499 AAGGAACTGTCAAGATGACCAATGACGACTTCCAGAGGCTCACCCGACGTACACGCC 558
 OY 186 AspIleLeuSerHisIleuHisIleuTyrLeuArgGlu 197
 Db 559 GACATCTCTCTCAGATCTGCACTACTGACAGAG 594

RESULT 13
 B0953977 1074 bp mRNA linear EST 21-AUG-2002
 LOCUS AGNCOURT 8877989 Lupski_sciatic_nerve Homo sapiens CDNA clone
 DEFINITION IMAGE:6198592 5', mRNA sequence.
 ACCESSION B0953977
 VERSION B0953977.1 GI:22369455
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1074)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: L1AM13610 row: h column: 17
 High quality sequence stop: 575.
 Location/Qualifiers
 1..1074
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="6198592"
 /clone_id="Lupski_sciatic_nerve"
 /sex="male"
 /tissue_type="sciatic nerve"

US-09-902-772-2 (1-451) x BM489636 (1-634)

```
QY      120  ThrAsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisVal 139
      |||
Db      3  ACCAATGACGAAGAGTTATTGTGTCAGACATCTTCAGTTATGTAGACACAGACCATGTA 62
QY      140  ArgGlnTrpLeuGlnTrpAlaValIleGluTyrGlyLeuProAspValAspIleLeuLeu 159
      |||
Db      63  CGGCGAGTGGCTGGAGTGGCGAGTGAAGAGATGCTTCAGACGTGACATCTGTG 122
QY      160  PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179
      |||
Db      123  TTCCAGAACATGTGATGGGAAGAGTGTGTAATAATGACCAAGATGACTTCAGAGACTC 182
QY      180  ThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluArg--- 198
      |||
Db      183  ACGCGAGCTATTAACGAGATATCTCTGTCACACCTACACTACAGAGAGACTCT 242
QY      198  ----- 198
Db      243  CTTCCACATTGTGACTTCAGATGATGTGATTAAGGCTTACAAAACCTCCACGGTTAATG 302
QY      199  -----GlyAlaThrPheIlePheProAsnThrSerValTyrProGlu 212
      |||
Db      303  CATGCTAGAAACACAGACAGAGCCACTTTATTTCCTCAATACATCAGTTTACCCAGAA 362
QY      213  AlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAla 232
      |||
Db      363  GCAACGCCAAGAATTAACACAAAGCCAGATTACCTATGAGCAAGCAGAGATCAGCG 422
QY      233  TrpThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrVal 252
      |||
Db      423  TGGACGAGTCAACAGCCATCCACTCACTCAAAAGCTACCAACCATCATCTTCAACAGTG 482
QY      253  ProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 272
      |||
Db      483  CCCAAAACAGAACACGACGCTCTCAGATTAGATCTTATCAGATTTTGGACGACGACG 542
QY      273  SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 292
      |||
Db      543  AGCGCTCTTGCAAATCCAGGAGTGGGAGATACAGCTATGGCAGTCTCCTACTGAGACTT 602
QY      293  LeuSerAspSerSerAsnSerAsnCysIle 302
      |||
Db      603  CTGTGGAGAGCTCCAACTCCAACTGCATC 632
```

Search completed: July 28, 2003, 07:39:12
Job time : 1906.27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:14:59 ; Search time 3810.11 Seconds

(without alignments)
3651.115 Million cell updates/sec

Title: US-09-902-772-4
Perfect score: 2588
Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPNTRLPAAHMPSHLCTYY 478

Scoring table: BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPFO.spool/US09902772/rnat_23072003_093657_14878/app_query.fasta_1.1294
-DB=GenBml -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09902772.GCgn_1_1_4707.ernat_23072003_093657_14878 -NCPV=6 -ICPV=3
-NO_MAP -LARGESUBSTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBml:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_inu:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	94.1	1516	5 GGERG	X77159 G.gallus ER
2	2436	94.1	1528	6 E31254	E31254 Protein hav
3	2293.5	88.6	2209	10 AB073080	AB073080 Mus muscu
4	2262.5	87.4	1413	5 AY065661	AY065661 Gallus ga
5	2262.5	87.4	1447	6 E31253	E31253 Protein hav
6	2186	84.5	2133	10 AB073079	AB073079 Mus muscu
7	2180.5	84.3	3166	9 HUMERG2	M17254 Human erg2
8	2180	84.2	1798	5 XLAJ4125	AJ224125 Xenopus l
9	2177.5	84.1	1808	10 AB031088	AB031088 Rattus no
10	2149.5	83.1	2131	10 AB073078	AB073078 Mus muscu
11	2015.5	77.9	2012	5 XLAJ4126	AJ224126 Xenopus l
12	1808.5	69.9	3126	9 HUMERG11	M21535 Human erg p
13	1491.5	57.6	1436	5 DRE249590	AJ249590 Danio rer
14	1482	57.3	1359	9 AY029368	AY029368 Homo sapi
15	1482	57.3	2394	9 BC010115	BC010115 Homo sapi
16	1482	57.3	2916	6 A36461	A36461 Sequence 2
17	1482	57.3	2938	6 AR080101	AR080101 Sequence
18	1482	57.3	2938	9 HSHOMFLI	X67001 H.sapiens H
19	1482	57.3	2957	9 HUMERGBFLI	M98833 Homo sapien
20	1482	57.3	2959	9 BC001670	BC001670 Homo sapi
21	1482	57.3	6849	12 AY029367	AY029367 Synthetic
22	1479	57.1	1673	9 S45205	S45205 Fli-1-Frien
23	1474	57.0	1729	10 MPEL11	X59421 Mouse Fli-1
24	1440.5	55.7	4403	5 XLEFLG	X66979 X.laewis MR
25	1432	55.3	3490	5 CCRNAFLI	Y14773 Coturnix co
26	1420	54.9	3545	5 CCFILONCO	Y14774 Coturnix co
27	1394	53.9	1932	9 HUMERFLIA	M93255 Human Fli-1
28	1341	51.8	1245	5 AE177538	AE177538 Danio rer
29	1253.5	48.4	816	10 S66169	S66169 Erg-3-lmmun
30	1036.5	40.1	762	9 S72621	S72621 EMS...erg f
31	834	32.2	1431	9 AF327066	AF327066 Homo sapi
32	809.5	31.3	14518	9 AP001423	AP001423 Homo sapi
33	809.5	31.3	100000	9 AP000021	AP000021 Homo sapi
34	809.5	31.3	100000	9 AP000163	AP000163 Homo sapi
35	809.5	31.3	340000	9 AP001731	AP001731 Homo sapi
36	778	30.1	588	9 S72865	S72865 EMS...EMS-e
37	702.5	27.1	66352	9 HSY17293	Y17293 Homo sapien
38	702.5	27.1	187131	2 AP001535	AP001535 Homo sapi
39	702.5	27.1	217382	2 AP001122	AP001122 Homo sapi
40	604	23.3	2180	3 AY060316	AY060316 Drosophila
41	602	23.3	634	3 SULENG	M61067 Sea urchin
42	514.5	19.9	491	3 HDI311813	AJ311813 Hediste d
43	485.5	18.8	169741	9 AC097468	AC097468 Homo sapi
44	484.5	18.7	83830	2 AC095777	AC095777 Rattus no
45	484.5	18.7	192968	2 AC127107	AC127107 Rattus no

RESULT 1

ALIGNMENTS

LOCUS	GGERG	1516 bp	mRNA	linear	VRT 27-APR-1995
DEFINITION	G.gallus ERG mRNA.				
ACCESSION	X77159				
VERSION	X77159.1				
KEYWORDS	ERG gene.				
SOURCE	Chicken.				
ORGANISM	Gallus gallus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;				
AUTHORS	Phasianine; Gallus.				
TITLE	1 (bases 1 to 1516) Dhordain, P., Dewitte, F., Desbians, X., Stehelin, D. and Mesodermal expression of the chicken erg gene associated with precartilaginous condensation and cartilage differentiation				
JOURNAL	Mech. Dev. 50 (1), 17-28 (1995)				
MEDLINE	95329425				
PUBMED	7605748				
REFERENCE	2 (bases 1 to 1516) Dutergue-Cogulland, M.				
AUTHORS	Direct Submission				
TITLE	Submitted (17-JAN-1994) M. Dutergue-Cogulland, CNRS UA 1160, Oncologie Moleculaire, Institut Pasteur, 1 rue Calmette, 59019 Lille, FRANCE				
JOURNAL					
FEATURES	Location/Qualifiers				
source	1..1516 /organism="Gallus gallus" /db_xref="taxon:9031" /tissue_type="spleen" /clone_lib="lambda gt10" /dev_stage="adult" 63..1499 /gene="ERG" 63..1499 /gene="ERG" /codon_start=1 /protein_id="CAA54404.1" /db_xref="GI:790440" /db_xref="SWISS-PROT:Q90837" /translation="MASTIKALSVSEEDSLFPCAYGSPHLAKTEMTASSSEYQOTD SKMSPRYQODMLISOPPARVATIKMECPNONGSRNSPDCEVAKGKMGSSDNGCM NIGYMEBKHKPPNMTNERRVATVPDPTLMSDHYRQMLEAVKREGLPDVILILE ONIDGKELCKMTKDDROLTPSYVADLLSHLYLRETPPLPHLSDDYKALQSPRL MHAINTGATPIPIPTNIVPEPATRITTRDPLPEQAKRSAMTSHSP70SKATOPSS STVKTEDQRLDPVQILGPTSSRLANPSGQIQLOMQLLELSDSNCSNCTIETEST NGEKMTPDEVARQMERKSKPMNMDKLSRALRYDYDKIMTKVGRKRAYFTDFH GIAALOPHPPESSMYKPSDLPYMSYVHAHPQKMNFAVBPALPTVSSSFAPAPNP YMNSTGCIYPTNTPALPAHMHSHGTY"				
BASE COUNT	458 a 392 c 327 g 339 t				
ORIGIN					
Alignment Scores:					
Spred. No.:	2.45e-158	Length:	1516		
Score:	2436.00	Matches:	458		
Percent Similarity:	91.97%	Conservative:	0		
Best Local Similarity:	91.97%	Mismatches:	0		
Query Match:	94.13%	Indels:	40		
	5	Gaps:	2		
US-09-902-772-4 (1-478) x GGERG (1-1516)					
Oy	1 MetlaserthrillysglualaleuserValyaserGluaspglnserleuePhaglu 20				
Db	63 ATGGCAAGACATTAAAGGAAGCACTTATCATGAGTGAAGCAACAGCTCTTTTGAG 122				
Oy	21 CysAlatYrgIySerProHsIleuAlaIysHicGluMetHrAlaSerSerSerGlu 40				
Db	123 TGTGCGACACGAGATGCCGCCACCTTGCAAGACAGAAATGACAGCGCTCTTCACAGGAA 182				
Oy	41 TYrtylglnThrSerlySerMetSerProArgValProGlnGlnasprPleuSerGlnPro 60				
Db	183 TATGGGCAAAATCAATGAAGATGAGCCCGCGGCTTCCACAGAGCATGTTATACACACC 242				

QY	61	ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn	80
Db	243	CCGGCCAGAGTTAACATTAAAGTGAAGGTAAACCCAAACACAGTTAAAGGGCAAGAT	302
QY	81	SerProAspAspCysSerValAlaLysGlyLysMetValSerSerAspAsnVal	100
Db	303	TCACCTGATGACTGCAGCGGTGGCAAAAGAGGAAATGTTAGCAGTTACAGCAATGTT	362
QY	101	GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr	120
Db	363	GGGAAGAACTAAGGAAGCTACATGGAAGAGCAATATTCGCCCTCCAAATATGCAACC	422
QY	121	AsnGluArgArgValIleValProAlaAspProThrIleTyrSerThrAspHisValArg	140
Db	423	AATGAACCAAGAGTTATTTGTGCCACAGATCCTACGTTATGGAGACAGACCATTTACGG	482
QY	141	GlnTrrPleuGluTrrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe	160
Db	483	CAGTGGCTGGAGTGGGCGACGTGAAGAGATAGTGTCTCCAGACGTGACATCTTGGTCTTC	542
QY	161	GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr	180
Db	543	CAGAACATTTAGGGGAAAGAGTTGTGTAAATGACCAAGATGATCTCCAGAGACTCCAG	602
QY	181	ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu	200
Db	603	CCGAGCTATTAACGCAGATATCTCTCTGTCACACCTACACTACCTCTGAGAGACACTCTCT	662
QY	201	ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis	220
Db	663	CCACATTTGACTCTCAGATGATGTGATGAAGCCTTAAACCTCCACGGTTAATGAT	722
QY	221	AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluVala	240
Db	723	GCTAATAAACACAGGAGGAGCACTTTATTTTCCAAATACATCAGTTACCAGGAACA	782
QY	241	ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrr	260
Db	783	ACGCAAAAMATTAACAACAGGCCAATTTACCTTATGAGCAAGCAGGAGATTCAGCGTGG	842
QY	261	ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro	280
Db	843	ACGATGCACAGGCATTCACACTCAGTCAAAAGCTACCCCAACCATCTCTTCAACAGTGGCC	902
QY	281	LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer	300
Db	903	AAAACAGAAGACCCACGCTCCCTCAGTTACATCTTTCAGATTTCTGGACCCAGCAGAC	962
QY	301	ArgLeuAlaAsnProGlySerGlyLysIleGlnLeuTrrPglIlePheLeuGlnLeuLeu	320
Db	963	CGTCTTGGAATCCAGGAGTGGGAGATACAGCTATGGCAGTTCTTAAGTGGAGCTCTTG	1022
QY	321	SerAspSerSerAsnSerAsnCysIleThrTrrPglGluTyrHisGlyGluIlePheLysMet	340
Db	1023	TCGGAACACTCCAACTCCAACTGCATCCCTGGAGGGCCACAATGGGGAATTCAGAATG	1082
QY	341	ThrAspProAspGluValAlaArgArgTrrPclGluArgLysSerLysProAsnMetAsn	360
Db	1083	ACAGACCTGTGAAGTGGCTCGGGCTTGGGGAGAGAGAAACCAACCTTAACATCAAC	1142
QY	361	TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal	380
Db	1143	TATGACAAACCTCAGCGCTGCTGCACTTGCTACTATGACAAATAATATTATGACTTAAGTT	1202
QY	380	-----	380
Db	1203	CATGGTAACGCTATGCTTACAAATTTGATTTCCACGGAATTCGCTCAGGGCCCTCCAGCT	1262
QY	381	HisProProGluSerSerMetLysTyrTrProSerAspLeuProTyrMetSerSerTyr	400
Db	1263	CACCTCCAGATCATCCATGTACAAATAACCATACAGACCTCCCTCATGATGATGTTCTTAC	1322

QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 Db 1322 ----- 1322
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 Db 1323 CATGACACACCCCAAGATGAATCTTGAGCTCCCAATCCCCCTGCTTGGCCGTAACC 1382
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPheSerProThrGlyIleTyr 460
 Db 1383 TCATCCAGCTTTTGGCTGCCCCCTATCATCTGGAATTCACCAATGGAGCATCTAC 1442
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 Db 1443 CCCAATACCAAGCTGCACTGCTCATATGCTTCCATCTTGGCACTACTAC 1496
 RESULT 2
 E31254 1528 bp DNA linear PAT 18-JUN-2001
 LOCUS E31254
 DEFINITION Protein having cell calcifying inhibitory activity and gene
 encoding the same.
 ACCESSION E31254
 VERSION E31254.1 GI:13025686
 KEYWORDS JP 199075871-A/2.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1528)
 AUTHORS Hiroyasu, I., Yoshinobu, H., Maritjo, P., Joel, R. and Helena, E.
 TITLE Protein having cell calcifying inhibitory activity and gene
 encoding the same
 JOURNAL Patent: JP 199075871-A 2 23-MAR-1999;
 CHUGAI PHARMACEUT CO LTD, HE TRUSTEES OF THE UNIVERSITY OF
 PENNSYLVANIA
 COMMENT OS Unidentified
 PN JP 199075871-A/2
 PD 23-MAR-1999
 PE 29-MAY-1998 JP 1998166076
 PR 18-JUN-1997 US 08/878177, 20-JUN-1997 US 60/050297 PI
 HIROYASU IWAMOTO, YOSHINOBU HIGUCHI, MARITJO PASHIPIKI, PT JOEL
 ROZENBROOD, E.
 PI HELENA E.
 PC C12N15/09, A61K48/00, C07K14/465, C07K16/18, C12Q1/68//A61K38/00,
 PC C12P21/02,
 PC C12P21/08, C12N15/00, A61K37/02
 CC Strandness: Single;
 CC topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..1528
 FT 1.1528
 FEATURES
 source Location/Qualifiers
 1..1528
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 464 a 395 c 330 g 339 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,48e-158 Length: 1528
 Score: 2436.00 Matches: 458
 Percent Similarity: 91.97% Conservative: 0
 Best Local Similarity: 91.97% Mismatches: 0
 Query Match: 94.13% Indels: 40
 Db: 6 Gaps: 2
 US-09-902-772-4 (1-478) x E31254 (1-1528)
 QY 1 MetaLaserThrIleLysGlnAlaLeuSerValIleSerGluAspGlnSerLeuPheGlu 20
 Db 63 ATGGACAGCATTATTAAGGAAGCATTTATCAGTGTGAGTGAAGACAGCTCTGTTGTGAG 122
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 Db 1203 CATGTAACGCTATGCTACAAATTTGATTTCCACGGAATGCTCAGGCCCTCCAGCCT 1262

Db 122 TGTGCTACGAGATCGCCGCCACCTTGCAGAAAGACAGAAATGACAGCCTCTTCCAGTGA 182
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValAlaProGlnGlnAspTyrLeuSerGlnPro 60
 Db 183 TATGGGCAAAACATCAAGATGAGCCCCCGCTTCCACAGAGAGCTGTTATCACAGCCCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 243 CCGGCCAGAGTTACCATTTAGATGAGTGTAAACCCAAACAGGTTATATGGTCAAGGAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyIleLysMetValSerSerSerAspAsnVal 100
 Db 303 TCACCTGATGATGCAGCGCTGGCAAAAGAGGGAATAATGTTAGACAGTTACACAAATGTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGlyLysHisIleProProAsnMetThrThr 120
 Db 363 GGGATGAACATATGAGAGCTACATGAGAGAGAGATATTCCTCCCAATATGACAAACC 422
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140
 Db 423 AATGAACGAGAGTTATTTGCTCCAGACAGATCTTACGATTTGAGACACAGATGACGG 482
 QY 141 GlnTyrLeuGluTyrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
 Db 483 CAGTGGCTGAGTGGGAGCTGAGAGGATGTGCTTCCAGAGCTGAGACATCTTGTGTTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheIleArgLeuThr 180
 Db 543 CAGAACATGATGAGGAAAGATGTGTAAAGACCAAGAGAGACTCCAGAGACTCAGG 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLysHisTyrLeuArgGluThrProLeu 200
 Db 603 CCGAGCTATACAGCAGATATCTCTCTGTACACCTATCACTACCTAGAGAGACTCTCTT 662
 QY 201 ProHisLeuThrSerAspAspValLysPylasAlaLeuGlnAsnSerProArgLeuMetHis 220
 Db 663 CCACATTTGACTTTCAGATGATGATGATTAAGGCTTACAAACCTCCACAGGTTATGAT 722
 QY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 Db 723 GCTAAGAACACAGGAGAGACCTTTATTTTCCAAATAACATCAGATTATCCACAAAGCA 782
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 260
 Db 783 ACGCAAGACATTAACAAAGCCAGATTTACTTTAGACAAAGCAGAGATCAGCGTGG 842
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValPro 280
 Db 843 ACGAGTCAACAGCATCCCATCCAGTCAAAAGCTACCAACCATCATCTTCAACAGTGGCC 902
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 Db 903 AAAACAGAGACACCGCTCTCAGTTAGATCTTATCAGATCTTGGAGCCGACAGCAC 962
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheGlnPheLeuLeuLeuLeu 320
 Db 963 CGTCTTGCAGAAATCCAGGAGATGGGACAGATACAGCATATGGCACTTCTCTAGAGCTTCTG 1022
 QY 321 SerAspSerSerAsnSerAsnGlyIleThrTyrGluGlyThrAsnGlyLeuPheLysMet 340
 Db 1023 TCGGACACCTCCAAATCCAACTGATCACTCGGAGGGCCAAATGGGAGATTCAAAGAT 1082
 QY 341 ThrAspProAspGluValAlaArgArgTyrPheGlyLysGlySerLysProAsnMetAsn 360
 Db 1083 ACAGACCTGATGAGAGTGTGCTGGCTGGGAGAGAGAAAGCAAACTTAACATGAGAC 1142
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 Db 1143 TATGCAAACTAGCGCGTGCACCTTGCCTACTATGACAAATAATATATATGCTAAAGTT 1202
 QY 380 ----- 380
 Db 1203 CATGTAACGCTATGCTACAAATTTGATTTCCACGGAATGCTCAGGCCCTCCAGCCT 1262

QY	381	HisProProGluSerSerMetEtyrLysTyPProSerAspIaeProTyrMetSerSerTyr	400
Db	1263	CACCTCCAGATATCATCATCATATACAAATACCATACAGACCTCCCTACATAGTTCCTAC	13222
QY	401	HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro	420
Db	1322	-----	13222
QY	421	HisAlaHisProGlnLysMetAspPheValAlaProHisProProAlaLeuProValThr	440
Db	1323	CATCCACACCCCCAGAGATGAATACCTTGTTACTCTCCCATCCCTCGCTTTCCCGTAACC	1382
QY	441	SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyr	460
Db	1383	TCATCCAGCTTTTGTGCTGCCCTTAATCCATACACGAAATTCACCAACACGAGGCATCTAC	1442
QY	461	ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr	478
Db	1443	CCCCAATACACAGCTGCACGAGCTGTCAATATGCTTCCCATCTTGGACACTACTAC	1496
RESULT 3			
AB073080			
LOCUS	AB073080	2209 bp	mRNA
DEFINITION	Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,		
ACCESSION	AB073080		
VERSION	AB073080.1	GI:16197544	
KEYWORDS			
SOURCE	Mus musculus		
ORGANISM	Mus musculus cDNA to mRNA, clone:Erg-3.		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.		
JOURNAL	Mus musculus Erg mRNA		
REFERENCE	Published Only in Database (2001)		
AUTHORS	2 (bases 1 to 2209)		
TITLE	Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan		
AUTHORS	1-7-22 Shohri-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)		
FEATURES			
SOURCE	Location/Qualifiers		
	1..2209		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="16"		
	/clone="Erg-3"		
	1..2209		
	/gene="Erg"		
	148..1608		
	/gene="Erg"		
	/note="homolog of human ets-related gene (ERG), transcript variant 3"		
	/codon_start=1		
	/protein_id="BAB69550.1"		
	/db_xref="GI:16197545"		
	/translation="MIDTYDPAKHIREALSYSEDSQSLDECAYGTPHLAKTEMTASS		
	SSDGGQSKSPRYPODDWLSQLARVTIMECPNPOVNSRNSPDSCSVNKGKMG		
	SPVTGNSYSYMEEKHPPNMNTNERRIVADPLTMSTDHYRQMLEAVVEYGLL		
	DVVLLEQNIIDKEIKELKMTKDDPQRLSYNADILSHLYLRTPHLNLSDDVDAKA		
	LONSRLMAHNTGGAAFIPTNVSVEAORTTRTDRLYEPPRSAGMGLHTPQ		
	SKAASPSAVPRTQEQRPQIDPXYQIGTPSSRLANGSQOIQMLGILBELSSNS		
	NCITWESTNEKFKTDEVARRRGEKSPNMNYSKLSALRYIDKNTIMTYVHGKR		
	YARKEPDHGIQAQLQPHAPRESSLYKYPQSDLYMGSYAHAPQKNFVSFPHPALPVTS		
	SRFAPSPYNNSPSGIYPTNRLPASHMPESHGTYT"		
BASE COUNT	606 a	608 c	545 g 450 t
ORIGIN			

Alignment Scores:			
Pred. No.:	2,34e-148	Length:	2209
Score:	2293.50	Matches:	431
Percent Similarity:	86.76%	Conservative:	15
Best Local Similarity:	86.55%	Mismatches:	11
Query Match:	86.62%	Indels:	41
DB:	10	Gaps:	3
US-09-902-772-4 (1-478) x AB073080 (1-2209)			
QY	2	AlaserthrileysglualaleuSERvalSergluaspglnSerleupheglucys	21
Db	172	GCAGCTCATATTAAAGAGAGCCCTTGCACTTGAGACGAGACCACTCATATTAGTGT	231
QY	22	AlatryglSerPROHisleuAlalysrthrglmettrAlaSerSerSergluytr	41
Db	232	GCCATAGGAAGCCACCTGGCTTAGACAGATGACCGCATCTTCCAGTACATAT	291
QY	42	GLylnlrrSerlysmetSerProarvalProGlnlAspTrpleuSerGlnPro	61
Db	292	GGCCAGACATCCAAATGATAGTCCGAGATCCTCAGCAGAGATGGCTGTCAACCCCA	351
QY	62	AlaArvalthrileysMetglucysAnproasnGlnAlaSnGlySerAragsnsr	81
Db	352	GCCAGGGGACCAATCAAGATGAGAGCAACCTTAGCTAGGTAAGTTCGAGAACTCA	411
QY	82	ProAspAspCysSerValAlaLalselyllysmetLalseserSerAspAsnValgly	101
Db	412	CCTATGTAGTCCACAGTGTACAAAGAGTGGAGATGGTGGCAGCCCGATCTGTGGG	471
QY	102	MetaSnryrGlySerTryMetGluGluYushHisIleProProCAsmetrthrAsn	121
Db	472	ATGAGCNRAGGCGACGTNATGAGAGAGAGATGTGGCGCTCCAAATATAGCCAAAT	531
QY	122	GLuarArvalIleValProAlaAspProthrleutrpSerThrAspHisValArgln	141
Db	532	GAGCGAGAGATGATCGTCCGAGATCTCTGTGTGAGACACAGACCATGTCCGACAG	591
QY	142	TrypleuGluTrpAlaVallysGlnTryGlyLeuProAspValAspIleleuLeuPheGln	161
Db	592	TGGCTGGAGTGGGGGTGAAGAAATATGGCTCCCTCGATGTGGAGAGCTTACTATTAG	651
QY	162	AsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro	181
Db	652	AATATCGATGGGAAGAGACTGTGCAGATGTACAAGAGATGACTTCCACGCGTCCAGCGG	711
QY	182	SerTryrAsnAlaAspIleleuLeuSerHisleuHisTryleuArgGlnThrProleuPro	201
Db	712	AGCATCAATGGCGACATCTTCTCCACATCTCCACTCCAGTCCAGAGACTCCCTTCCA	771
QY	202	HisleuThrSerAspAspValAspLysAlaleuGlnAsnSerProArgLeuMetHisAla	221
Db	772	CATCTGACTTCCGATGAGCTGTATAGAGCTTTACAAACTCTCCACGGTTAATGATGCC	831
QY	222	ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTryProGlnAlaThr	241
Db	832	AGAAACACAGGGGGGCGAGCTTTATTTATTTCCAAATACTTCAGATATCCGAAGCTACG	891
QY	242	GlnArgIleThrThrArgProAspLeuProTryGluGlnAlaArgArgSerAlaThrPhe	261
Db	892	CMAAGAAATTACACTAGGCCAGATTACTTATGAGCCCTCCAGAGATCAGCCGAGAC	951
QY	262	SerHisSerHisPProThr--GlnSerLysAlaThrGlnProSerSerSerThrValPro	280
Db	952	GGCCACAGCCACCTCACCCCTCAGTCCAAAGCTGCTAGCCATCTCCTTGCACTGCC	1011
QY	281	LysThrGlnAspGlnArgProGlnLeuAspProTryGlnIleleuGlyProThrSerSer	300
Db	1012	AAACCTGAGAGCCAGCGCTCTCACTTATGATCTTACACGATTCGTGGAGCGACAGTAGC	1071
QY	301	ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeuLeu	320

Db 1072 CGCCTTCTATCCAGGAGTAGTGGCCAGATCCAGCTGTGGAGCTTCCTGCTGCAACTCCTG 1131
Qy 321 SerAspSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 340
Db 1132 TCACACAGCTCCACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1191
Qy 341 ThrAspProAspGluValAlaArgArgTrrpGluArgTrrpGluArgTrrpGluArgTrrpGlu 360
Db 1192 ACACACCCGAGCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1251
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTrrpGluArgTrrpGluArgTrrpGluArgTrrpGlu 380
Db 1252 TATGACAGCTCAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311
Qy 381 HisProProGluSerSerMetTrrpLysTrrpProSerAspLeuProTrrpMetSerSerTrr 400
Db 1311 ----- 1311
Qy 401 HisGlyLysArgTrrpAlaTrrpLysPheAspPheHisGlyLeuAlaGlnAlaLeuGlnPro 420
Db 1312 CACGGAGAGCGCTACGGCTACAGTTGACTCCACGGAGTGGCCAGGCCCTGCAAGCC 1371
Qy 420 ----- 420
Db 1372 CACCCCTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1431
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1432 CACGCCACCCCGAGAGATGAACTTGTGCTCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1491
Qy 441 SerSerSerPhePheAlaAlaProAsnProTrrpAsnSerProThrGlyGlyLeuTyr 460
Db 1492 TCTTCCAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db 1552 CCGAACACTAGGCTCCAGCAGCATATGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1605

RESULT 4
AY065661
LOCUS 1413 bp mRNA linear VRT 17-DEC-2001
DEFINITION Gallus gallus erg isoform C-1-1 mRNA, complete cds; alternatively
spliced.
ACCESSION AY065661
VERSION AY065661
KEYWORDS AY065661.1 GI:17887440
SOURCE Gallus gallus.
ORGANISM Gallus gallus.
REFERENCE 1 (bases 1 to 1413)
AUTHORS Iwamoto,M., Higuchi,Y., Koyama,E., Enomoto-Iwamoto,M., Kurisu,K.,
TITLE Yeh,H., Abrams,W.R., Rosenbloom,J. and Pacifici,M.
JOURNAL Transcription factor ERG variants and functional diversification of
MEDLINE chondrocytes during limb long bone development
PUBMED J. Cell Biol. 150 (1), 27-40 (2000)
20351415
10893254
2 (bases 1 to 1413)
REFERENCE Iwamoto,M., Higuchi,Y., Enomoto-Iwamoto,M., Kurisu,K., Koyama,E.,
AUTHORS Yeh,H., Rosenbloom,J. and Pacifici,M.
TITLE The role of ERG (ets related gene) in cartilage development
JOURNAL Osteoarthritis. Cartil. 9 Suppl A, S41-S47 (2001)
21535378
11680687
3 (bases 1 to 1413)
REFERENCE Iwamoto,M., Higuchi,Y., Yeh,H. and Pacifici,M.
AUTHORS Direct Submission
TITLE Submitted (05-DEC-2001) Oral Anatomy and Developmental Biology,
JOURNAL Osaka University Faculty of Dentistry, 1-8 Yamadaoka, Suita, Osaka
565-0871, Japan
FEATURES Location/Qualifiers

source 1..1413
/organism="Gallus gallus"
/db_xref="taxon:9031"
37..1392
CDS
/note="alternatively spliced; transcription factor"
/codon_start=1
/product="erg isoform C-1-1"
/protein_id="FA140889.1"
/db_xref="GI:17887441"
/translation="MASTIKELASVSDOSLEPCAYGSPHLAKTEMTASSSEYGO
T SKSPRYPODDWLSQPPARVITKMECPNOVNGSRNSPDDCSVAKGKRWSSDNVGM
NYSYMEKEHIPPENMTNERRVYIPADPTLMSYDHYRQLEMAVEYGLPDVILF
ONTDEKELCKMTKDDFORLPSYNAIDILSHLYRREGAFETIPNTSVYBEAORT
TRDPLPYEORARSSAMWSHSPOTCAKOPSSPYKTEDORPOLDPYQIIGPSSRLA
NPGSGQIQLOMFLLELSDSNSCTTMEGTNGEFKATDDEVARRKERKSNMAY
DKLSRALRYFYDKNIMTKVHKGRAYRFDHGLAQALQPPPESSMYKYPDSLPYMS
YHAHPQKMFVAPPPALPVTSSEFAAPNPYMSPTGIGYIPNTRPALAHMPSHLGTY
Y"

BASE COUNT 424 a 370 c 311 g 308 t
ORIGIN

Alignment Scores:
Pred. No.: 1,82e-146 Length: 1413
Score: 2262.50 Matches: 430
Percent Similarity: 86.35% Conservative: 0
Best local Similarity: 86.35% Mismatches: 1
Query Match: 87.42% Indels: 67
DB: Gaps: 3

US-09-902-772-4 (1-478) x AY065661 (1-1413)

Qy 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 37 ATGGCAAGCACTATTAAAGAGCACTTATCAGTGTGAGTGAAGACAGCTCTGTTGAG 96
Qy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
Db 97 TGTGCTTACGAGTGGCCCACTGCAAGACAGAAATGACAGCTCTCTCCAGAGAA 156
Qy 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrrpLeuSerGlnPro 60
Db 157 TATGGCGAACAATCAAGAGTACGCCCGGCTGCCAGCAGGACTGTTATCACACCCC 216
Qy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 217 CCGGCCAGATTACCATTAAGAGAGTGTACCCAAACAGGTTAATGGGTCAAGGAAT 276
Qy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 277 TCACCTGATGCTGCACGCGTGGCAAAAGAGGAAATGGTTACACTTCAGACAAATGTT 336
Qy 101 GlyMetAsnTrrpGlySerTrrpMetGluGlnLysHisIleProProProAsnMetThrThr 120
Db 337 GGGATGACATATGGAACCTCATGGAAGAGACATATTCGCCCTCCAAATATGACAAAC 396
Qy 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrrpSerThrAspHisValArg 140
Db 397 AATGACAGAGAGTTATGTGCGCAGCATCTCATGATGAGACAGACATGTCAGG 456
Qy 141 GlnTrrpLeuGlnTrrpAlaValLysGluTrrpGlyLeuProAspValAspIleLeuPhe 160
Db 457 CAGTGGCTGAGAGTGGCAGAGTGAAGAGATATGCTTCCAGACGTCGACATCTGTTTC 516
Qy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 517 CAGAACATTCATGGGAAGAGTGTGTGTAATGACCAAAATGACTCCAGAGACTCAGC 576
Qy 181 ProSerTrrpAsnAlaAspIleLeuLeuSerHisLeuHisTrrpLeuArgValTrrpLeu 200
Db 577 CCGAGCTATACGACAGATATCTCTCTCACCTTACACTTACACAGAGAGA----- 630
Qy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220


```

Db      603 CCGAGCTATACGAGATATCTCTGACACCTACACTACCTCAGAGAGA----- 656
Oy      201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnInsSerProAlaGluMetHis 220
Db      656 ----- 656
Oy      221 AlaArgAsnThrGlyAlaThrPheProAsnThrSerValTyrProGluAla 240
Db      657 -----GGAGCCACTTTTATTTTCCAAATCATCATGTTTACCCAGAGCA 701
Oy      241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnAlaArgArgSerAlaTyr 260
Db      702 ACGCAAGAAATTAACAACAGCCAGATTACCTATAGAGCAAGAGATCAGCGTGG 761
Oy      261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      762 ACGAGTACAGCCATCCACTCAGTCAAAAGCTAACCCATCATCTTCAACAGTCCC 821
Oy      281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      822 AAAACAGACAGCCAGCGCTCTCCAGTTAGATTCATAGATCTTGGACCGACGACGAC 881
Oy      301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuGluLeu 320
Db      882 CGTCTTCAATCCAGGAGAGTGGGACATACAGCTATGACGAGTCTTCTGAGACTTCTG 941
Oy      321 SerAspSerSerAsnSerAsnCysIleThrThrProGlnGlyThrAsnGlyLupHelyMet 340
Db      942 TCGGACAGCTCAACTCACTCATCATCCTGGGAGGAGCAAAATGGGAGTTCAAGATG 1001
Oy      341 ThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsn 360
Db      1002 ACGACCTGATGAAAGGCGCTGGGGAGAGAGGAAACCAACCTAATCATGATAC 1061
Oy      361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db      1062 TATGACAACTCAGCCGCTGACTGCTACTACTATGACAAATAATATTATGACTAAGATT 1121
Oy      380 ----- 380
Db      1122 CATGGTAAACGCTATGCCCTCAAAATTTGATTTCCACGGAATCGCTCAGCCCTCAGCCT 1181
Oy      381 HisProGluSerSerMetLysTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1182 CACCCCTCAGATATCATGATGATACAAATACCAATACCACTCCCTTCAATGAGTTCCTAC 1241
Oy      401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1241 ----- 1241
Oy      421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1242 CATGCAACCCCGAGAGATGAATCTTGTAGCTCCATCCCTCGCTTGGCCGCTAAC 1301
Oy      441 SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyLysIleTyr 460
Db      1302 TCATCCAGCTTTTGTGTCGCCCTAATCCATACGGAATTCACCAATCGAGGAGATCTAC 1361
Oy      461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1362 CCCAATACCAAGGCTGCGACGCTGCATATGCTTCCATCTTGGCACCTACTAC 1415

```

RESULT 6
 LOCUS AB073079 2133 bp mRNA linear ROD 17-OCT-2001
 DEFINITION Mus musculus Ery mRNA, mouse homolog of human ets-related gene ERG,
 transcript variant 2, complete cds.
 ACCESSION AB073079
 VERSION AB073079.1 GI:16197542
 KEYWORDS Mus musculus cDNA to mRNA, clone: Ery-2.
 SOURCE Mus musculus
 ORGANISM

```

REFERENCE
AUTHORS Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.
TITLE Mus musculus Ery mRNA
JOURNAL Published only in Database (2001)
REFERENCE 2 (bases 1 to 2133)
AUTHORS Ozawa, R., Noguchi, H., Taylor, T. D., Takeda, T., Hattori, M. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
FEATURES
source
1. 2133
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
/clone="Ery-2"
1..2133
/gene="Erg"
149..1540
/feature="Erg"
/note="homolog of human ets-related gene (ERG), transcript variant 2"
/codon_start=1
/protein_id="BAB6949.1"
/db_xref="GI:16197543"
/translation="MIQVPPPAHIEKALVSSEDSLEFCAATPPLAKTEMTASS
SDVGRTSKSPRYPOODWLSQAPARVITIKKCPNPOVNSRNSGKMG
SPDVMSYCSYMEKRPBPNTNRRTVLPADPLMSDHYRDLKPAVREYGL
DVALEONIDKELCKMTKDDRLRPLPSNADILLHLRLRTPPLPLSDVDKA
LQNSPRLMHRNRTGGAATFEPNTSVYEAQRIITRAADQSPSAVETEDQRLDP
YQIIGPTSRLANPNSGOIQIOWFLLELSDSSNSITWGTGFERKMDPDEARR
WGERKSPKNNMYDLISRALRYDKNTVHGKRYAVKFFPHGIAQALDPHPRESSL
YKYPDLPYMGSYVAHPQKMNKFPHPALPVTSSSFASNPWNPTGSIYNTRL
PASHMPSHLCTYV"
BASE COUNT 587 a 579 c 531 g 436 t
ORIGIN
Alignment Scores:
Pred. No.: 5,34e-141 Length: 2133
Score: 2186.00 Matches: 413
Percent Similarity: 85.31% Conservative: 11
Best Local Similarity: 83.10% Mismatches: 11
Query Match: 84.47% Indels: 62
Gaps: 3
US-09-902-772-4 (1-478) x AB073079 (1-2133)
Oy      2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db      173 GCACCTCATATTAAGAGGCGCTTGTCACTGTGTGAGCGACGACGATATTGAGATGT 232
Oy      22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41
Db      233 GCCTACGGAAGCCACACCTGGCTAAGACAGAGTGAACCGCATCTTCCAGAGTAT 292
Oy      42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnProPro 61
Db      293 GCCAGACATCCAGAGATGCTCCAGAGTCCCTTGAAGAGGAGTGGCTGTCAAGCCCCA 352
Oy      62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db      353 GCCAGGCTCAACATCAAGATGAGGTGCAACCTTGTGAGTGAATGGTTCCAGCAACTCA 412
Oy      82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db      413 CTCATGATGATGACAGTGTGAACAAAGTGGGAGAGTGTGGGACCCGAGATACTGTGGG 472

```

OY	102	MetAnuYrGLySerTyrMetGluGluGluYrHisSiliEProProPAsmMethrThrAsn	121
Db	473	ATGACCTAGCGACGCTACTATGAGGAGAAAGCATGTGCGCTCCCAATATGACCAAAAT	532
OY	122	GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln	141
Db	533	GAGCCGACAGATGATGTCCTCCATCATCTACTCTGTGGAGCACAGACATGTCCGACAG	592
OY	142	TrpLeuGluTrpAlaValIleGluGluYrGlyLeuProAspValAspIleLeuLeuPheGln	161
Db	593	TGGCTGGAGTGGCGGTGAAAGAAATATGGCCCTCCGATGTGAGAGCTTACTATATTTGAG	652
OY	162	AsnIleAspGlyIleGluLeuLeuCysLysMethTrpLysAspAspPheGlnArgLeuThrPro	181
Db	653	AATATCGATGGGAAGAGAGCTGTGCAGAAATGACAAAGATGATTCACAGCGCTACGCGC	712
OY	182	SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro	201
Db	713	AGCTTCAAATGGCGACATTCCTCTCCACATCTCCACTCCACTCCAGAGAGACTCCCTTCCA	772
OY	202	HisLeuThrSerAspAspValAspLysAlaLeuGlnAnSerProArgLeuMethHisAla	221
Db	773	CATCTGACTTCGCGATGAGCTGTATATAGGCTTTACAAACTCTCCACGGTTAATGATGCC	832
OY	222	ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr	241
Db	833	AGAAACACAGGGGGGCGACAGCTTTATTTTCCAAATACTTCAGTATATCCGCAACGTACG	892
OY	242	GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerLalaThrThr	261
Db	893	CAAAGAAATTACACTAGGCCA-----	913
OY	262	SerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValProLys	281
Db	914	-----GCTGCTCAGCATCTCCCTCTGCAGATGCCAA	946
OY	282	ThrGluAspGlnArgProGluLeuAspProTyrGlnIleLeuGlyProThrSerSerArg	301
Db	947	ACTGAGACCCAGCGCTCTCAGTATATCTTTACCGAATCTGGGACGACGATGTAGCGCG	1006
OY	302	LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSer	321
Db	1007	CTTGCTATTCAGAGTAGTGCCAGATTCAGCTGTGGAGTTCCTGCTCGAACTCCTGCA	1066
OY	322	AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGlnPheLysMethThr	341
Db	1067	GACACCTCCAACTCCAACTGATCATCCCTGGGAAGCACCAACGGGAGTTCAAGTTGACA	1126
OY	342	AspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsmMethAsnLys	361
Db	1127	GACCCGGAGAGGTGTGCTCGCGCTGGGGGGGAGAGGAAGACGACCAACATGTGAACTAT	1186
OY	362	AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMethThrLysValHis	381
Db	1187	GACAAAGCTCACCGCGCCCTCCGCTACATACGACAAACATCATGACCAAGAGTG---	1243
OY	382	ProProGluSerSerMetTyrLysTrpProSerAspLeuProTyrMetSerSerTyrHis	401
Db	1244	-----CAC	1248
OY	402	GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro---	420
Db	1247	GGGAAGCGCTACGGCTACAAAGTTTGACTTCACCGGGAATGGCCAGGCCCTCAGGCCCA	1306
OY	421	-----His	421
Db	1307	CCTCCAGTGCCTCTGTACAGTACCCCTCCGACCTGCCATACATGAGGCTCCTATAC	1366
OY	422	AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer	441
Db	1367	GCCCAACCCCAAGAAATGATCTTGTGTGCTCCCAACCTCCCGCTCTCCAGTCACTACT	1428
OY	442	SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro	461

Db	1427	TCGAGTTCTTGGTCTTCCCGCAGCCACACTGGAATTCACGAGTGGGGCATCTACCCG	1486
Qy	462	AsntrhargleuproAlaAlaHisMetProSerHisIleuGlyThrYrYr 478	
Db	1487	AACACTAGGCTCCGACGACGCATATGCTCTCACCCTGGGACACTACTAC	1537
RESULT 7			
HUMERG2			
LOCUS	HUMERG2	3166 bp ss-RNA	linear PRI 08-NOV-1994
DEFINITION	Human erg2 gene encoding erg2 protein, complete cds.		
ACCESSION	M17254		
VERSION	M17254.1	GI:182186	
KEYWORDS	erg 2 protein.		
SOURCE	Human cell line COLO 320, CDNA to mRNA, clone lambda 12.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Rao,V.N., Pappas,T.S. and Reddy,E.S.		
TITLE	erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation and translation		
JOURNAL	Science, 237 (4815), 635-639 (1987)		
MEDLINE	87263429		
PUBMED	3299708		
FEATURES			
source	Location/Qualifiers		
	1..3166		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="21q22.3"		
	257..1645		
gene	/gene="ERG"		
CDS	257..1645		
	/gene="ERG"		
	/note="erg 2 protein"		
	/codon_start=1		
	/protein_id="AA452398.1"		
	/db_xref="GI:182187"		
	/db_xref="GBD:G00-119-884"		
	/translation="MIQTVDPAAHIKEALVSVSDLSFECAVGTPLHAKTEMTASS		
	SSDIDGNTSKMSRPVQDDWLSSOPARVYTIMECNPSQVNSSRNSPDECSVAKGKMG		
	SPDVGKSGKSYMEKEKMPNMTINERATVVPADPLVSTDRVQMLEAVKYGCLP		
	DVNILLFONIDGKEIKMTKDDFORLLPSVNAIDILSHLYLRETPLPHLISPDVKA		
	LQNSPRLMARANDLPEYPPRRSAWCHGPQSKAQPSPSTVPEORQDQPL		
	QILPTPSRLANPSSGOIQLMQFLLELSLSSNSSCTTGWCTNGERKMTDPDVAWR		
	GERSKRNMTDKLSRALRYDYKNIIMTKVHGKRAYKFDPHGIAALDPHPRESSLY		
	KYSPDLPEMGSYNAHPQKNFVAPHPALPVTSSFFAAPNPYMNPSPTGQIYNTRLP		
	TSNHPSHGTRY"		
	291..298		
repeat_region	/note="8-bp repeat sequence"		
repeat_region	509..516		
polya_signal	/note="8-bp repeat sequence"		
	3143..3149		
	/note="putative"		
BASE COUNT	227 bp upstream of PvuII site; chromosome 21.		
ALIGNMENT	915 a 770 c 768 g 713 t		
ALIGNMENT SCORES			
Pred. No.:	2.06e-140	Length:	3166
Score:	2180.50	Matches:	413
Percent Similarity:	84.74%	Conservative:	9
Best Local Similarity:	82.93%	Mismatches:	11
Query Match:	84.25%	Indels:	65
DB:	9	Gaps:	4
US-09-902-772-4 (1-478) x HUMERG2 (1-3166)			
Qy	2	AlaserthriletysgualalaleuSerValserGluaspGlnserleupheglucys 21	
Db	281	GCACCTCATATCAAGAGCGCTTATCGTTGTGAGTAGAGACCACTGCTTTGACTGT 340	
Qy	22	AlatryglserProHisleuAlaIalysThrGluMetThrAlaSerSerSergIuYr 41	

Db 341 GCCTACGGAGAACCCACCTGGCTAAGACAGATGACCCGGCTCCTCCACGAGACTAT 400
 GY GlycylthrSerIysMetSerProArgValProGlnGlnAspTyrPheLeuSerGlnProPro 61
 Db 401 GGGAGAGCTTCCAAAGATGAGCCACGCGCTCCCTCAGAGAGATGGCTGCTCAACCCCA 460
 GY AlaArgValThrIleIysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 Db 461 GCCAGGGTCACCATCAAAATGATGTAACCTTACCCAGAGTGATGGCTCAAGAACTCT 520
 GY ProAspAspCysSerValAlaAlaGlyGlyIleIysMetValSerSerSerAsnValGly 101
 Db 521 CCTGATGATGACATGTGGCCAAAGCGGAGAGATGGTGGGCGACCCACAGACACCGTGGG 580
 GY MetSerIyrGlySerIyrMetGluGluIleHisIleProProProAsnMetThrThrAsn 121
 Db 581 ATGAGACTACGGCAGCTCATGAGAGAGAGACATGACACCCCAACATGACACAGAAC 640
 GY GluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArgGln 141
 Db 641 GAGGCGAGAGTATCGTCCGACGAGATCCCTACGATGAGATGACAGACATGTCGGCAG 700
 GY TrpLeuGluTyrPalaValIleValGlyIleIleLeuProAspValAlaIleLeuPheGln 161
 Db 701 TGCGTGGAGTGGCGGTGAAGAATATGGCTCCACAGCTCAACATCTTGTATTCAG 760
 GY AsnIleAspGlyIysGluLeuGluCysLeuMetThrIyrAspAspPheGlnArgLeuThrPro 181
 Db 761 AACATCAATGAGGAGAAAGATGTGCAAGATGACCAAGAGACGCTCCAGAGCTCACCCC 820
 GY SerIyrAsnAlaAspIleLeuLeuSerHisIleuHisIleuArgGluThrProLeuPro 201
 Db 821 AGCTACAAAGCGGACATCTCTCTCATCTCCACATCTCCACAGAGACCTCTTCCA 880
 GY HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
 Db 881 CATTTGACTTCAGATGATGTGATTAAGCTTACAAACCTCCACAGGTTAATGACAGCT 940
 GY ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
 Db 941 AGAAGACACA----- 949
 GY GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaThrThr 261
 Db 950 -----GATTTACCATATAGACCCCGCCAGAAATCAGCTTGAGCC 988
 GY SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 Db 989 GGTACAGGCGCACCCACGCCCCAGTGCAGAAAGCTGCTCAACCATCTCTCCACAGTGC 1048
 GY LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 Db 1049 AAAACCTAAGACCGCTCTCTCACTTATGATTCCTTATGAGACCAACAAAGTACG 1108
 GY ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGluLeuLeu 320
 Db 1109 CGCCTTGCMAATCCAGGACAGTGGCCACATCCAGCTTTGGCAGTTCTCTCGAGCTCTCG 1168
 GY SerAspSerSerAsnSerAsnCysIleThrTyrPgluGlnIleThrAsnGlyIlePheLysMet 340
 Db 1169 TCGGACAGCTCCAACTCCAGTGCATCAGCTGGGAGAGCAACCAAGGAGTTCAAGATG 1228
 GY ThrAspProAspGluValAlaAlaArgArgTyrGlyIleArgLysSerLysProAsnMetAsn 360
 Db 1229 ACGGATCCCGACGAGGGTGGCGGCTGGGAGAGGAGGAGCAACCAACATCAAC 1288
 GY TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrIleVal 380
 Db 1289 TAGCATAGCTCAGACCGCGCTCTCTACTACTATACAGAAACATCATGACCAAGCTC 1348
 GY HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 Db 1348 ----- 1348

GY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 Db 1349 CATGGAGAGCGCTACGCTTACATGATGACTTCCACGAGATGCCAGGCTTCAGGCC 1408
 GY 420 ----- 420
 Db 1409 CACCCCGGAGTATCTGTACAGTACCCCTCAGAGACCTCCGATAGTGGCTCAT 1468
 GY HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 Db 1469 CACGCCACACACAGAGATGACTTGTGGCGGCCACCTCCAGCCCTCCCGTGACA 1528
 GY 441 SerSerSerPhePheAlaAlaProAsnProTyrThrAsnSerProThrGlyIleTyr 460
 Db 1529 TCTTCCAGTATTTTGTGCTGCCCAACCCCATCTGAGAAATTCACCAACTGGGGGATATAC 1588
 GY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
 Db 1589 CCAGAACCTAGGCTCCCGACAGCATATAGCTTCTCATGTGGCGACTTACTAC 1642
 RESULT 8
 XLAJ4125 1798 bp mRNA linear VRT 21-JAN-2000
 LOCUS XENOPUS LAEYIS ERG GENE (ERG_A).
 DEFINITION A1224125
 ACCESSION A1224125
 VERSION A1224125.1 GI:5420045
 KEYWORDS ERG gene; transcription factor.
 SOURCE Xenopus laevis.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 1798)
 AUTHORS Baltzinger M., Mager-Heckel A.M. and Remy P.
 TITLE Xl erg: expression pattern and overexpression during development
 JOURNAL Dev. dyn. 216 (4-5), 420-433 (1999)
 MEDLINE 2009678
 PUBMED 10633861
 REFERENCE 2 (bases 1 to 1798)
 AUTHORS Baltzinger M.
 TITLE Direct Submission
 JOURNAL Submitted (24-FEB-1998) Baltzinger M., UPR9005, MMCD, Centre
 National de la Recherche Scientifique, 15, rue Rene Descartes,
 67084, FRANCE
 FEATURES
 source location/Qualifiers
 1. 1798
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="lambda ZAPIT-erg A"
 /clone_id="lambda ZAPIT"
 /dev_stage="stage 30 embryo"
 321..1778
 /gene="erg"
 321..1778
 /gene="erg"
 /function="transcription factor"
 /codon_start=1
 /protein_id="CAB46566.1"
 /db_xref="GI:5420046"
 /db_xref="SPTREMBL:O9W700"
 /translation="MOHMAFTPTVEDYSNEALSVSSEDSGLFECTYGTPLHTYTEMTA
 SSSSDYGOTSRKMSPRVQODLSQPSRVTIKMCNPSYNGSRSSPDDCSIGKSKM
 GGSDVNGMYGSEYIEKHISPMNTNERRVIVPADPTLMSDHYHOMLEMAIKGYG
 LPDVVLLFONIDGKELKMTKEFHLTPSYNADILSHLHVLRETPPLPLTSDVD
 KALONSRLMHARNTGASPIFPNSSVYODANORIRPSRODLSYSPSRSAWTHHPAP
 SKASQSTYPRKPEDRPQDIPQIIGTSSRLANSGOIOIMOLLELSSSSNSN
 CITWEGNGEFKMTDPEVARRMGERSKRNMTDKSLRLRYTDNKTIVHGKRY
 AYKFDHGLAQLQPHPESTMYKYSELPYSSYHAAHPQKNMFVAPHPALPVTSSS
 FFAAPNAYNMSPGSITPNTRLPASHMSHLGTY"

BASE COUNT
 ORIGIN

561 a 429 c 358 g 450 t

CDs

63. .1430
/gene="Vespl4"
/note="similar to human erg1"
/codon_start=1
/product="vascular endothelial cell specific protein 14"
/protein_id="BAB62744.1"
/db_xref="GI:15128489"
/translation="MASTIKEALSVSKDOSLFRCAYGTPLHAKTEMTASSSDYQGT
SKMSPRVQODMLSQPARVITKMECPNSOYNGSRNPDECSYTKGKMGSPITVGH
SYGIMEKMKHPNMTNERRVIVPADPLWSTDHVRQLEMAVKEKGLIDVLE
ONIDKELCKMTKDFOFLTPSYNADILSHLHYLRETPPLHLSDDVAKLQMSPL
MHANPDLPEPPRPSTWTGSHPTPOSKAAPSPSTVPTEDORPOLDPYOLIGPIS
SLRANPSGOLOLWOFLELSDSSNSKTEWTEGPFKMDDEVARMRGKESRP
NMNYDKLSRLRYDDKIMTKVHKRYAFKDFDHGIALQAPHPRESSLYKPSDLP
YMSYTHHPKMFNVAPHPALPTPTSSFPATPMPNPSPTGIGTIPNTRLPASHMPSH
Lgity."

BASE COUNT 510 a 522 c 426 g 350 t

ORIGIN

Alignment Scores:

Pred. No.: 1.67e-140 Length: 1808
Score: 2177.50 Matches: 412
Percent Similarity: 84.37% Conservative: 9
Best Local Similarity: 82.57% Mismatches: 13
Query Match: 84.14% Indels: 65
DB: 10 Gaps: 4

US-09-902-772-4 (1-478) x AB031088 (1-1808)

QY 1 MetaIaSerThrIleTylsGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 63 ATGGCCAGCACTATTAGAGAGCCTTGCTGCTTGAGCAAGACAGCAGTCACTTTGAG 122
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB 123 TGTGCTACAGGAGACGACACCTGCTAGAGCAAGAAATGACCGCATCTCTCCAGTGAC 182
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPheSerGlnPro 60
DB 183 TATGCCAGACATCCAGATGAGTCCAGAGTCCCAACAGAGAGTGGCTCTCTCAGGCC 242
QY 61 ProAlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 243 CCAGCCAGGCTCACCATCAAGATGAGTGCACACCTTGCAGCGTGAAGGTTCCAGAGAC 302
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB 303 TCACCTGACGATGACAGTGTACCAAGGTGGGAAGATGTGGGACGCTGATACGTG 362
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB 363 GGAATGAGCTACGCGACGCTACATGAGAGACACATGCGGCCCCCAATATACACACG 422
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTyrSerThrAspHisValArg 140
DB 423 AATGACGACAGAGTATGCTCCGAGATCTCTGCTGAGACACAGACAGTCTCGG 482
QY 141 GlnTyrPheGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB 483 CAGTGGCTGAGTGGGAGTGAAGAAATACGCGCTCTAGACGTCGCTTATATT 542
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB 543 CAGATATGACGGGAGAGAGCTGTGCAAGATGCAAAAGTGACTTCCACAGGCTCACT 602
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB 603 CCGAGGTACATGCGCATCTCTCTCAGATCTCCACATACCTCAGAGAGACCTCCCT 662
QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB 663 CCACATCTGACATCCGACGAGCTGTGATTAAGCTTTACAAAACCTCCACAGGTTATCAT 722

QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
DB 723 GCTAGAAACACA----- 734
QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnAlaAlaArgSerAlaTrp 260
DB 735 -----GATTACTTATGAGCCTCCACAGAGATCAACCTGG 770
QY 261 ThrSerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerThrVal 279
DB 771 ACCGGCCACACCCACCCACCCCTAGTCCAGCAAGCTGCCACCATCTCCCTCCACACTG 830
QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB 831 CCCAAATGTGAAGACACACGCTCTGATTAATACCTTACCAAGATCTGGGACCCACACGT 890
QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeu 319
DB 891 AGTCGCTTGCTAATCCAGTATGAGGCCAGATCCAGCTGTGGAGTTCTCTGAGAACTC 950
QY 320 LeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheLys 339
DB 951 CTGTCTGACAGCTCCACTCCACTCCACTGCTGAGGAGGACACCAACGGGAGATTCAAG 1010
QY 340 MethrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 359
DB 1011 ATGACAGACCCGGATGAGGTGCTGCGGGGGGAGCGGAAGACCAACCCCAACATG 1070
QY 360 AsnTyrAspLysLysSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB 1071 AACATGACAAACACAGCGCTGCGCTCCGCTACAGTACAGCAAAACATATGACCAAG 1130
QY 380 ValHisProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSer 399
DB 1131 GTG----- 1133
QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB 1134 ---CACGGGAAGCGCTATGCTTACCAAGTTTACTTCCACGSGAGATTGCCAGGCCCTG 1190
QY 420 Pro----- 420
DB 1191 CCCCATCCCCCGAGTGTGCTGCTGACAAAGTACCCCTCCAGCTGCCGTATAGGGCTCC 1250
QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
DB 1251 TATCACACCCACCCCGAAGATGAACTTGTGCTCCGCCCTCCGCGCTCCAGTTC 1310
QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle 459
DB 1311 ACATCTTCCAGTTCTTGTGCTACCCCGAACCACATCTGGAATTTGCCGACTGGGGGCATC 1370
QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
DB 1371 TACCGAAGACTAGAGTCCACGACGACCATATGCTCCCTCCACGCGGAGACTACTAC 1427

RESULT 10
AB073078 2131 bp mRNA linear ROD 17-Oct-2001
LOCUS AB073078
DEFINITION Mus musculus Erg mRNA, mouse homolog of Human ets-related gene ERG,
transcript variant 1, complete cds.
ACCESSION AB073078
VERSION AB073078.1 GI:16191716
KEYWORDS
SOURCE
ORGANISM Mus musculus cDNA to mRNA, clone:Erg-1.
REFERENCE 1. Mus musculus
Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murine: Mus.
AUTHORS Osawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and
Sakaki, Y.
TITLE Mus musculus Erg mRNA
JOURNAL Published only in Database (2001)

AUTHORS	2 (bases 1 to 2131)					
TITLE	Ozawa, R., Noguchi, H., Taylor, T.D., Takeda, T., Hattori, M. and Sakaki, Y.					
JOURNAL	Direct Submission Submitted (15-OCT-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suhiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) Location/Qualifiers					
FEATURES	Source					
gene	1..2131 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="16" /clone="Erg-1" 1..2131 /gene="Erg" 148..1536 /gene="Erg" /note="homolog of human ets-related gene (ERG), transcript variant 1"					
CDS	/codon_start=1 /protein_id="BAB69948.1" /db_xref="GI:16191717" /translation="MIGTVPDAAHKEALSVSDESLECAVGTPLHAKTEMTASSSSDGGTSKMSPRVPOQDWLSQAARVTIKMECPSONGSRNSPDDCSVKRGKMVGSPDVGMGSYSGYMEKKHVPPENMTNRRRVYPADPTLMSTDHYROWLFWAVKEYGILLDVDVLIFONIDGKELECKMKTKDFQLRLPSYNADLLSLHLRTPLPHLSVDVKAALONSPLRHMANHTDLPEYEPRSASWTGHSHLTPOSKAOPSVAVEPDOPQLDLPDYIOLNPSTSRLANPGSGOLOLOMPLEILLSOSNCNTJWETCNEFEFTMDPYAPRWGERSKRMNMYDKLSRALRYTYDKNINIMTKHGKRYATKFPDHGLAOALOHPHPRESSILYKIPDELPMGSTYHAHPQKMNVSHPPALPVTSSEFSAISPNIWNSPFGITYPNTRLPASHPSHLGYST"					
BASE COUNT	582 a 590 c 530 g 429 t					
ORIGIN						
Alignment Scores:						
Pred. No.:	1.71e-138 Length: 2131					
Score:	2149.50 Matches: 408					
Percent Similarity:	84.14% Conservative: 11					
Best local Similarity:	81.93% Mismatches: 14					
Query Match:	83.06% Indels: 65					
DB:	10 Gaps: 4					
US-09-902-772-4 (1-478) x AB073078 (1-2131)						
Oy	2 A laserthrilllelysgluAlaleuSeXValSerGlunaspGlnSerleuePhleglucys 21					
Dd	172 GCACCTCATATTGAAGAGCGCCCTTGTCAGTTGTGAGCGAGCACGACTCATTTGAGAGT 231					
Oy	22 AlAtyrGlSerProHisLeuAlAllysThrGluMetThrAlaserSerSerSerGluTy 41					
Dd	232 GCCAACGGAAACGCACACCCTGGCTPAAGACAAGATGACCCTTCCTTCAGAGACTAT 291					
Oy	42 GlyInThrSerLyMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61					
Dd	292 GGCCAGACATCCAAAGATGAGGCCAGAGTCCCTCAGCAGAGCTGGCTGTCCAAGCCCCA 351					
Oy	62 AlaArgValThrllelysmetGluCysAsnProasnGlnValasnglySerArgAsnser 81					
Dd	352 GCCAGGGGCACCAATCAAGATGAGATGCAACCTTAGTACAGGAATGGATTCCAGAGACTCA 411					
Oy	82 ProAspAspCySerValAlAlalelyGlyLySmetValSerSerSerAspAsnValgly 101					
Dd	412 CCTGATGAGTCACTGTGCACAAAGGTGGAGAAGTGGTGGCCACCCGGATACTGTGGGG 471					
Oy	102 MetAnfnyrGlySerTyrmGluGluLylusHisIleProProAsmMetThrThrAsn 121					
Dd	472 ATGAGCTACGGCAGCTCATGATGAGAGAGACATGTGCCGCTCCCAATATGACCAAAAT 531					
Oy	122 GlunrgrgvalllilealproAlaAspprothrleutrpserthrAsphlsvalArgln 141					
Dd	532 GAGGCGAGAGATGATGTCCTGCGAGATCTCACTGTGGAGCACAGACATGTCCGACAG 591					

OY	14	TripLeuLurPrAlaValIysGluYrGlyLeuProAspValAspIleLeuLeuPheGln	161
Db	592	TGCGTGAGGCGGCGGTGAAGAAATATGCGCTCCCTCGATGtGAGAGTtTACTTTCAG	651
OY	162	AsnIleAspIlyGluLeuGlySerIleuLeuSerThrIySaSaSpPheGlnArgLeuThrPro	181
Db	652	AATATCATGATGAAGAGCGGTGCAAGATGACAAAGATGATGACTTCACCGGCTCACGCCG	711
OY	182	SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluThrProLeuPro	201
Db	712	AGCTACAAATCCCAACATTTCTCTCCATCTCCACTCCATCCAGAGAACTCCCTTCCA	771
OY	202	HisLeuThrSerAspAspValAspIlySaIaLeuGlnAsnSerProArgLeuMetHisAla	221
Db	772	CATCTGACTTCGCAATGACGTTGATAAAGCTTTACAAACTCTCCACGGTTAATGATGCC	831
OY	222	ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr	241
Db	832	AGAAACACA-----	840
OY	242	GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaThrThr	261
Db	841	-----GATTTACCTTATGAGCCTCCACGAGAGTACGCTTGAGC	879
OY	262	SerHisSerHisProThr---GlnSerIySaIaThrGlnProSerSerSerThrValPro	280
Db	880	GGCCACAGACCACTTACCCCTCAGTCCAAAGCTGTACAGCATCTCCCTCGAGTGC	939
OY	281	LysThrGluAspGlnArgProGluLeuAspProTyrGlnIleLeuGlyProThrSerSer	300
Db	940	AAACTGTAAACACGAGCTCCTAGTTAGATCCTTACAGATCTCGGACCGACCAAGTAC	999
OY	301	ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglIleLeuLeuGlnLeuLeu	320
Db	1000	CGCCTTCCTAATCCAGATGATGCGCAGATCCAGCTGTGAGCTTCGCTCGAACTCTG	1051
OY	321	SerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGluPheIyMet	340
Db	1060	TCAGACAGCTCCAACTCCAACTGCATCACTGGGAAGCAACCAAGGGGATTCAAATG	1111
OY	341	ThrAspProAspGluValAlaArgArgTrpGlyGluArgIySerIyProAsnMetAsn	360
Db	1120	ACAGACCCGGACAGAGTGGCTGCGCGCTGGGGGAGAGGAAGCAAGCAATCATGAA	1171
OY	361	TyrAspIySerIySerIyArgIaLeuArgTyrTyrTyrAspIyAsnIleMetThrIyVal	380
Db	1180	TATGACAAAGTGCAGCGCGCCCTCGCTACTATCTACGACAAACATCATGACCAAGTG	1231
OY	381	HisProProGluSerSerMetIyIyTyrProSerAspLeuProTyrMetSerSerTyr	400
Db	1239	-----	123
OY	401	HisGlyIyAsArgTyrAlaTyrIyIyPheAsPheHisGlyIleAlaGlnAlaLeuGlnPro	420
Db	1240	CACGGGAAGCGCTACCGCTTCAAAAGTTTGACTTCACAGGGATTTGCCACGGCCCTG	1291
OY	420	-----	420
Db	1300	CACCCCTCGAGTCGCTCCCTGTACAAGTACCCCTCCGACCTGCATATACAGGCTCTAT	1351
OY	421	HisAlaHisProGlnIyMetAsnPheValAlaProHisProProAlaLeuProValThr	440
Db	1360	CACGCCCAACCCCAAGAAAGATTTGTTGTCCTCCACCTCCCGCTCTCCAGTACA	1411
OY	441	SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyr	460
Db	1420	TCCTTCAGATTTCTTCTCTCCCGCAACCATATCTGGAATTCACGACATGGGGGCACTTAC	1471
OY	461	ProAsnThrArgLeuProAlaIleHisMetProSerHisIleuGlyIyThrTyrTyr	478
Db	1480	CGAACACTAGGCTCCACGACCAATATGACCTCTCACCTGGGCACTTACTAC	1533

RESULT 11
XLAJ4126 2012 bp mRNA linear VRT 21-JAN-2000
LOCUS
DEFINITION Xenopus laevis erg gene (erg_E).
ACCESSION AJ224126
VERSION AJ224126.1 GI:5420047
KEYWORDS ERG gene; transcription factor.
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 2012)
Baltzinger, M., Mager-Heckel, A. M. and Remy, P.
Xl erg: expression pattern and overexpression during development
pled for a role in endothelial cell differentiation
Dev. Dyn. 216 (4-5), 420-433 (1999)
JOURNAL MEDLINE
PUBMED 20099678
10633861
2 (bases 1 to 2012)
Baltzinger, M.
Direct Submission
Submitted (24-FEB-1998) Baltzinger M., UPR9005, NMDCD, Centre
National de la Recherche Scientifique, 15, RUE Rene Descartes,
67084, FRANCE
FEATURES
SOURCE location/qualifiers
1..2012
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="lambda ZAPII-erg E"
/clone_lib="lambda ZAPII"
/dev_stage="stage 30 embryo"
225..1595
/gene="erg"
225..1595
/gene="erg"
/function="transcription factor"
/codon_start=1
/protein_id="CABA6567.1"
/db_xref="GI:5420048"
/db_xref="SPRMBL:Q9W629"
/translation="MIQTPDPSAHKEALSVSEDSLECTYGTPLHRTKTEWTASS
SDVGGTSMKSPRYPOODMLOPSRVTIKMECPNSOVNGSRSSPDSCISGSKMG
GSDVGNNGSYIEKHISPNMTNRRVYVPADPTLMSDHYRDLNNAIKRYGP
DYVYLFRONIDGKELCKMKEDPRRLTPSTNADLLSHLRLKRGASFEPNSVTO
DANORIPSRKDLSTEPSRSAMWNHAPPSKASQSTVPTEDPRQDLDPYQILGPT
SSRLANGSQIOLMOPLELSDSNSNCITWEGTNGEFTMDPDEVARWGERKSK
PNMYDKLSRLALRYDYDKNIMTKVHGKRYAKFPDHFGLAQLOPHPPESTMYKPSSEL
PYMSYHAHPQKMFVAHPHPALPVTSSEFFAANAWNSPTGSIYENTRLPASHMS
HLGTRYV"
BASE COUNT 634 a 479 c 395 g 504 t
ORIGIN
Alignment Scores:
Pred. No.: 2.5e-129 Length: 2012
Score: 2015.50 Matches: 383
Percent Similarity: 81.38% Conservative: 19
Best Local Similarity: 77.53% Mismatches: 23
Query Match: 77.88% Indels: 69
DB: 5 Gaps: 5
US-09-902-772-4 (1-478) x XLAJ4126 (1-2012)
QY 5 lileysglualaleuServalvalSerGlunspglInserleupheglucysAlatyrGly 24
Db 258 ATCAAAAGAGCTCTTTCAGTGGTGAAGATCAGTCGTATTCGAGTGACCTACGGA 317
QY 25 SerprohstleualalysThrGlumethThralaserSerSerSerGluTyrGlyGlnThr 44
Db 318 AGCGCACATCTTACTAAACGAGATACCGCATCTTCCTCACTGACTAAGGCAACC 377
QY 45 SerlysmetSerProAlaGvalProGlnGlnAspTrleuSerGlnProProAlaArgVal 64
|||||

Db 378 TCGAAATGAGCTCCACGCTGCTCCACAGACAGTGGCTCCACAGCTCCATCCAGGCTG 437
QY 65 ThrilleysmetGlucysAsnProAsnGlnValnsnglySerAArgasnerProAsp 84
Db 438 ACCATCAAGATGATGATGACMACCTTACATCAATGATCAAGAGAGTCCCGGATGAC 497
QY 85 CysSerValAlalysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyr 104
Db 498 TGCAGTATAGGAAAGAT 557
QY 105 GlySerYrmetGlulglulysHisIleProProProAsnmetThrThrAsnGluArg 124
Db 558 GCGACGATACATAGAAAGAAACATTTCTCCCAACATGACACCAACCAATGACAGAGA 617
QY 125 ValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGlnTrpLeuGlu 144
Db 618 GTGATTTAGCTGCTGACCTTCTTGTGGAGCAGATCATGTGCACAGTGCCTGAG 677
QY 145 TrpAlaVallysgulTyrGlyLeuProAspValAspIleleuLeuPheGlnAsnIleasp 164
Db 678 TGGCAATATAAGATATAGGCTTCCAGATGTCATGCTGCTCCAAACATTCGAC 737
QY 165 GlylysgluLeuCylysmetThrLysAspAspPheGlnArgLeuThrProSerTyrAsn 184
Db 738 GCGAAGGATATATGACAGATGACCAAGAGATTTCCATGACGACGCGCAAGCTACAT 797
QY 185 AlaAspIleleuLeuSerHisLeuHisTyrLeuArgGluThrProleuProHisLeuThr 204
Db 798 GCTATATCTCTCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 839
QY 205 SeraspValAspLysAlaLeuGlnAsnSerProArgleuMetHisAlaArgAsnThr 224
Db 839 ----- 839
QY 225 GlylyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIle 244
Db 840 ---GGTGGCTTTTATTTTCCCAACATCGTTTATCAAGCGCAACCAAGATC 896
QY 245 ThrThrArgProAspLeuProTyrGlyGlnAlaArgArgSerAlaTrpThrSerHisSer 264
Db 897 CCCGACGCGAAGATTTATCATATGAGACCCCAAGAGAGATCTGTCGACAAATCATCT 956
QY 265 HisProThrGlnSerLysAlaThrGlnProSerSerSerThrValProlyThrGluAsp 284
Db 957 GCACCA---CCATCAAAAGATCTCAACCA---CTTACCAAGATTCCTCAAAACGAAAGAC 1010
QY 285 GlnArgProGlnLeuAspProTyrGlnIleleuGlyProThrSerSerArgLeuAlaAsn 304
Db 1011 CCCAGGCGACAGCTAGATCTTATCAGATTTCTTGGCCACACAGCCGACCTTGCAT 1070
QY 305 ProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuSerAspSerSer 324
Db 1071 CCAGGAGTGGGAGATTCAGCTCTGCGCAATTTCTTACTGCAATGCTTGGAGATGTTCC 1130
QY 325 AsnSerAsnCylyIleThrTrpGluGlyThrAsnGlyGluPheLysMetThrAspProAsp 344
Db 1131 AACTCCACATGCTACTTCTGGGAGAGCAACATGGAATTTAAGATGACCGATCCGAT 1190
QY 345 GluValAlaArgArgTyrGlyGluArgLysSerLysProAsnmetAsnTyrAspLysLeu 364
Db 1191 GAAGTGCACAGCTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250
QY 365 SerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisProProGlu 384
Db 1251 AGCGGTGACCTGCTTACTACTAGATGATGATGATGATGATGATGATGATGATGATG 1298
QY 385 SerSerMetLysTyrTrpProSerAspLeuProTyrMetSerSerTyrHisGlyLysArg 404
Db 1299 -----CATGGCAACGGC 1310
QY 405 TyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro----- 420
Db 1311 TATGCTACAAATTTGATTTTTCATGAGGATGCTCAAGCTCTTCAACTCACTCTCCAGAA 1370

Db 1218 GGGGGTATATACCCAGACTAGGCTCCACACGACATATGCTTCTCATCTGGCACT 1277

QY 477 TTTTyr 478
 |||||
 Db 1278 TACTAC 1283

RESULT 13
 DRE249590

LOCUS DRE249590 1436 bp mRNA linear VRT 03-MAR-2000

DEFINITION Danio rerio mRNA for Fil-1 protein (fil-1 gene).

ACCESSION AJ249590

VERSION AJ249590.1 GI:6006476

KEYWORDS fil-1 gene; Fil-1 protein.

SOURCE Danio rerio.

ORGANISM Danio rerio.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 1436)
 Brown, L.A., Rodaway, A.R., Schilling, T.F., Jowett, T., Ingham, P.W.,
 Patient, R.K. and Sharrocks, A.D.
 Insights into early vasculogenesis revealed by expression of the
 Ets-domain transcription factor Fil-1 in wild-type and mutant
 zebrafish embryos
 Mech. Dev. 90 (2), 237-252 (2000)

REFERENCE JOURNAL MEDLINE 20108585
 PUBMED 10640707

2 (bases 1 to 1436)
 Sharrocks, A.D.
 Direct Submission
 Submitted (29-SEP-1999) Sharrocks A.D., School of Biological
 Sciences, University of Manchester, Oxford Road, Manchester, M13,
 9PT, UNITED KINGDOM

FEATURES
 source
 1. Location/Qualifiers
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /dev_stage="20-24hr embryo"
 45.1400
 /gene="fil-1"
 45.1400
 /gene="fil-1"
 /function="transcription factor"
 /codon_start=1
 /evidence="experimental"
 /product="Fil-1 protein"
 /protein_id="CAB56832.1"
 /db_xref="GI:6006477"
 /translation="MDGTKEALSVSESDSLPEPPYAAAPLPKTDMTASGTODGQ
 THKINPIPOQEMINOPVRYNKRREYDTHNGSRSPDCSGCNKMGVGTASQMY
 TGYMDEKCAPPMWTRERVYTPADPSLMSPDHYKRWMLAIKEGLDITTRHRS
 TDGERLCKMSKIDFLRLTSYNTVEVLISHLNYLRESSISITNTPSHADSPLAAAD
 DASYDAVRRRTGWSNMHSGSPVVSQSNPDQRPDPDYOLIGTPSSRLANG
 SGOIQLOMQLLELSDSANAGCITWEGTNEFKMTDPEVARRMGEKSKPMNYDL
 SRALRYDYDNIMTKYGRKRAYKPDFHGTAOLQPHPEVSTMYKYPSELYPYSVHA
 HOOKVNFVSHPPSMPTVSSNFRGPTTPYVSSPPGGIYPIPVSPRIANSHVPSHLSY
 Y"

BASE COUNT 377 a 439 c 349 g 271 t

ORIGIN

Alignment Scores:
 Pred. No.: 1 51e-93 Length: 1436
 Score: 1491.50 Matches: 289
 Percent Similarity: 68.13% Conservative: 53
 Best Local Similarity: 57.57% Mismatches: 85
 Query Match: 57.63% Indels: 75
 DB: 5 Gaps: 7

US-09-902-772-4 (1-478) x DRE249590 (1-1436)

QY 1 MetaLaserThrilletySguAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 |||
 |||||
 Db 45 ATGACGCAACTATTAGGAGCGGCTGTCACTGTGTGATGTAAGACCACTCTCTGTTGAG 104

QY 21 CysAlaLayr---GlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
 |||
 |||
 Db 105 CTTCCATACGCCCGCTGCTCTCTTACCCAAAGACAGATGCTGGCGGACGAC 164

QY 40 GluTyrGluInPheSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGln 59
 ::::|
 |||
 Db 165 GATTACGGCCAGACGCAAAATATCAACCCATTCCCTTACGAAAGGTGATTATACG 224

QY 60 ProPheAlaArgValThrIleLysMetGluCysAsnProAsnGlnValaAsnGlySerArg 79
 |||
 |||
 Db 225 ---CCGGTGGGGGAGACGTCAGACGAGAGAT-----GACCACATCAACGATCCAGA 275

QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
 |||||
 |||||
 Db 276 GAGTCCGCCGGTGAAGCTCCACCGGGGAAATGCAATTAATATGTTGGCGGAGGAGCGC 335

QY 100 ValGlyMetAsnTyrGlySerTyrMetGluGluTyrHisIleProProAsnMetThr 119
 |||||
 |||||
 Db 336 TTCACATGTAACATTAATCTACTGCTACATGATGAGAAATGTCTCCGCTCCAAACATACG 395

QY 120 ThrAsnGluArgArgValIleValProAlaAspProThrLeuTyrPheSerThrAspHisVal 139
 |||||
 |||||
 Db 396 ACCAATGAGAGAGATCATGCTCCCGACACCCGCTCTGTGGTCTCCGATACAGTG 455

QY 140 ArgGluTyrPheGluTyrPheValValLysGluTyrGlyLeuProAspValAlaPheLeuLeu 159
 |||||
 |||||
 Db 456 CGGACGTGTAAAGCTGGCCATTAAAGATACGCTTTCAGAGATGACACGCGCAT 515

QY 160 PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179
 |||
 |||
 Db 516 TTTCACACGACAGATGAGAAAGACTCTGCAAGATGAGCAAAAGCACTTCGAGACTC 575

QY 180 ThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrPro 199
 |||
 |||
 Db 576 ACCAGCCTTATTAACACGAGGCTCTCTCATCATCAATTAACCTCAGGAAAGT--- 632

QY 200 LeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProHisGluMet 219
 632 ----- 632

QY 220 HisAlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGlu 239
 |||||
 |||||
 Db 633 -----AGCTCATGATATCATACAAACAGCATCTCACGCTGAC 671

QY 240 AlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAla 259
 ::|
 |||
 Db 672 CAGTCCCGCTGCGCTGCTGCAAAAAGACGATGCGTATGATGCTGACGAGAGACGGC 731

QY 260 TyrThrSerIleSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrVal 279
 |||
 |||
 Db 732 TGGTCGAACAACATGCACACCGGAAGGCTCTCCACATGCTCTTCAGACGCTGCC 791

QY 280 ProLysThrGluAspGlnAAspProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
 |||||
 |||||
 Db 792 AAAAACCCGATCAGCTCTGCTCTGACGACGATCTTATGATATTAGAACCCACACG 851

QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheLeuLeuGluLeu 319
 |||||
 |||||
 Db 852 AGTCGACTTCTTAACACGAGCTCAGTCAGATCCAGCTGAGGAGTCTCTGTCGAGACTC 911

QY 320 LeuSerAspSerSerAsnSerAsnGlyIleThrTyrPheGluGlyThrAsnGlyLysPheLys 339
 ::|
 |||
 Db 912 CTGTCCACAGCGCCCAACGCGGCTGATACCTGGGAAGAACCAACGAGAGTTCACG 971

QY 340 MetThrAspProAspGluValAlaArgArgTyrGlyLysSerLysProAsnMet 359
 |||||
 |||||
 Db 972 ATGACGACCCACACAGAGTGGCAGCGCTGGGGCGCAACGCAACCAACACATG 1031

QY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
 |||||
 |||||
 Db 1032 AACCTAGACAAAGCTGACGCCGCTGCTACTACTACAGCAAAAGAACATCATGACCAAA 1091


```

|||||
Db 805 ACCAGAGTCGCTAGCCACCTGGAGAGGGGAGATCCAGCTGTGGCAATTCCTCTG 864
Qy 318 GtluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGlu 337
Db 865 GACTGTCTCCACAGCCGACCCAGCCAGCTGTATCCCTGGAGGGGAGCAACAGGGGAG 924
Qy 338 PhylSerMetThrAspProAspGluValAlaArgTrpGlyGluArgGlySerSerPro 357
Db 925 TTCAAAATGACGAGACCCCGATGAGTGGCCAGCGGCGGCGAGGAAACCAAGCC 984
Qy 358 AsnMetAsnTyrAspTyrLeuSerAspGalaLeuArgTyrTyrAspTyrAsnIleMet 377
Db 985 AACATGATATGACAGCAAGTACGACCGGCGCTCCGTTATATCATATATAAAACATTATG 1044
Qy 378 ThrValHisProProGluSerSerMetTyrTyrProSerAspLeuProTyrMet 397
Db 1045 ACCAAAGTG----- 1053
Qy 398 SerSerTyrHisGlySerArgTyrAlaTyrTyrPheAspPheHisGlyIleAlaGlnAla 417
Db 1054 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCCACAGGCAATTGCCAGGCT 1104
Qy 418 LeuGlnPro----- 420
Db 1105 CTCGACGACACATCCGACCGAGTGTGCATGTACAGTACCTTCTGACATCTCTACATG 1164
Qy 421 -----HisAlaHisProGlnTyrMetAsnPheValAlaProHisProProAlaLeu 437
Db 1165 CTTTCTCATTACGATGCCCGACGAGAGGTGACATTGTCTCCCTCCATCATCTCATG 1224
Qy 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGly 457
Db 1225 CCGTGCCTTCCTCCACCTCTTGGAGCGCATCAAAATGAGGAGCCCTCCCGAGGG 1284
Qy 458 GlyIleTyrProAsn-----ThrArgLeuProAlaHisMetProSerHisLeu 474
Db 1285 GGAATCTACCCCAACCCCAACGTCGCCCGCATCTTACACCAACGATGCTTACACTTA 1344
Qy 475 GlyThrTyrTyr 478
Db 1345 GGCAGCTACTAC 1356

RESULT 15
BC010115 2394 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:19589 IMAGE:3635042, mRNA, complete cds.
DEFINITION BC010115
VERSION BC010115.1 GI:14603315
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2394)
Strausburg, R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhr1.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,

```

```

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Mastello, C., Mastriani, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantliop, S., Thomas, P.J.,
Tlonsoson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.T., Walker, M.A.,
Zhang, L., H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 27 Row: 9 Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 257353.
FEATURES
source
1. 2394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:19589 IMAGE:3635042"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH-MGC.8"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
126. 1484
/codon_start=1
/product="Unknown (protein for MGC:19589)"
/protein_id="AAH10115.1"
/db_xref="GI:14603316"
/translaction="MDGTIKREALSVSDQSLFSDAYGAAILPKADMTASGSPDYGO
PKINILPQDEINOPVRVNYREYEDHNGSSESVDSVSCSLVGGSESPNMY
NSYMDKNGPPEPMTNERVYLPADPTLMTGEHMLKYSLSMEIDTSPFO
NMGRKELCKKMKREDPLRATYLYTEVLISLVSRLSLAYWTSTHTOOSRLSYKE
DPSEYDVRGACNMNMNSGLNKSPGLGNOTISKMPORPOPPYIILPTSRILNP
GSGQIOLKQPLLELSDSANASCTITEGNGEKMTDPPEVARMERKSRPMNTDK
LSRALRYVDKNTMTKRVKRAVKRFDPEHIALDLPPESSMYRYPEDISITMPSYH
AHQKXNFVPPHPSMPVSSSFFGAASOYWTSPITGIYPNPVPRNPTVPSHLS
Y"
BASE COUNT 669 a 609 c 574 g 542 t
ORIGIN
Alignment Scores:
Pred. No.: 1.27e-92 Length: 2394
Score: 1482.00 Matches: 291
Percent Similarity: 68.65% Conservative: 55
Best Local Similarity: 57.74% Mismatches: 80
Query Match: 57.26% Indels: 78
DB: 9 Gaps: 9
US-09-902-772-4 (1-478) x BC010115 (1-2394)
Qy 1 MetAlaSerThrIleGlyGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20
Db 126 ATGGACGGGACTATTTAAGAGGCTGTGCTGATGACACACACAGCTCCCTTTGAC 185
Qy 21 CysAlaTyrGly--SerProHisLeuAlaValThrGluMetThrAlaSerSerSer 39
Db 186 TCAGCGTACGAGCGGAGCGGACCCATCTCCCAAGCGGACATGATGCTCGGGAGATCCT 245
Qy 40 GluTyrGlyGlnThrSerTyrMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
Db 246 GACTACGGGAGCGCCACCAAGATCAACCCCTCCACACAGCAGAGATGATCAATCAG 305
Qy 60 ProProAlaArgValThrIleGlyMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 306 --CCAGTGGAGGTCAACGTCACGCGGAGTAT-----GACCACATGAATGATCCAG 356
Qy 80 AsnSerProAspAspCysSerValAlaValGlyGlyMetValSerSerSerAspAsn 99
Db 357 GACTCTCGGTGATGACGAGCGTTAGCAATGACAGCAAGCTGTGGCGGAGCGAGTC 416
Qy 100 ValGlyMetAsnTyrGlySerTyrMetGluGluValHis---IleProProProAsnMet 118
Db 417 AACCCCATGATACATCAACAGCTATATGAGACGAGAAATGAGCCCTCTCCCAACATG 476
Qy 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138

```

Db 477 ACCACCAAGAGAGAGATCATGTCCTCCGACAGCCACACATGTCAGACAGAGCAT 536
QY 139 ValArgGlnThrPheGlnValValysGlnThrLeuProAspValAspIleLeu 158
Db 537 GTGAGGCAATGGCTGGAGTGGCCATAAGAGATACAGCTTGATGAGATGACACATCC 596
QY 159 LeuPheGlnAsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArg 178
Db 597 TTTTCCAGAAATGATGACAGAAAGAACTGTGTAATGAACAAGAGAGAACTTCCTCCGC 656
QY 179 LeuThrProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlnThr 198
Db 657 GCCACCACTCTCAACACGGAAGTGTGTGCACACCTCAGTACCTCAGGGAAGT 716
QY 199 ProLeuProHisLeuThrSerAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
Db 717 TCACTG----- 722
QY 219 MethIAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
Db 723 -----CTGGCTATATACACCTCCACACC 749
QY 239 GlnAlaThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSer 258
Db 750 GACCAATCTTCACGATGAGTCAAGAAGACCTTCTTATGACTCAGTCAGAAAGAGA 809
QY 259 AlaThrThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
Db 810 GCTTGGGGCAATACATGAAATTCGGCCTCACAACAAGTCTCCTCCCTGGAGGGGCACA 869
QY 278 ThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
Db 870 ACGATCAGTAAGATACAGCAAGCAAGCCCGCCAGCATCCGTATCAGATCCCTGGCCCG 929
QY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeu 317
Db 930 ACCAGCAGTCGCTAGCCAACTGGAAGGGGAGATCCAGCTGTGCAATTCCTCTG 989
QY 318 GlnLeuLeuSerAspSerSerAsnSerAsnGlyIleThrThrPheGlnGlyThrAsnGlyLys 337
Db 990 GAGCTGCTCCGACAGCGCAAGCCAGCTGTATCAGCTGGAGGGGACCAAGGGGAG 1049
QY 338 PheLysMetThrAspProAspGlnValAlaArgArgThrPheGlnArgLysSerLysPro 357
Db 1050 TTCAAATGACGAGCCCGATGAGTGCCAGGCGCTGGGCGAGCGAAAGCAAGCCC 1109
QY 358 AsnMetAspTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
Db 1110 AACATGAAATACGACAGCTGAGCGGGCCCTCGTTATTAATTAATAAACATTATG 1169
QY 378 ThrLysValHisProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397
Db 1170 ACCAAAGTG----- 1178
QY 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
Db 1179 -----CACGGCAAAAGATATGCTTACAAATTGATTCACGGCATTGCCAGGCT 1229
QY 418 LeuGlnPro----- 420
Db 1230 CTGAGCCACATCCGACGAGTCGTCATGATACCAAGTACCCCTTGTGACATCTCTACATG 1289
QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
Db 1290 CTTTCCATCATGGCCACAGCAGAAAGTGAATTTCTCCCTCCCATCTCATG 1349
QY 438 ProValThrSerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGly 457
Db 1350 CTTGTCACTTCTTCACTTCTTGGAGCGCATCAATATGACCTCCCGACGGGG 1409
QY 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
|||||

Db 1410 GGAATCTACCCCAACCCCAAGTCCCGCCATCTTACACCAAGTGCCTTCACTTA 1469
QY 475 GlyThrTyrTyr 478
Db 1470 GGCAGCTACTAC 1481

Search completed: July 28, 2003, 06:34:30
Job time : 3830.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:05:39 ; Search time 306.661 Seconds

(without alignments)
3510.246 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKKALSVSEDSQSLFE.....IYPTRLRPAHMSHCTYY 478

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPto.spool/US09902772/runtat_23072003_093656_14868/app-query.fasta_1.1294
-DB=N_Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum2 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09902772_EGCGN1.1_364_etunat_23072003_093656_14868 -NCPu=6 -ICPu=3
-NO_MMAP -LARGESQDBRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.101002:*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	94.1	1528	20	AAAX26552
2	2262.5	87.4	1447	20	AAAX26551
3	1482	57.3	2938	14	AAO50644
4	1482	57.3	2937	24	ABK84139
5	1464.5	56.6	2954	22	AAH02915
6	1370	52.9	1890	14	AAO50662
7	801	31.0	567	22	ABA48124
8	801	31.0	567	22	ABA66003
9	801	31.0	567	22	ABA33090
10	801	31.0	567	22	AAK14426
11	801	31.0	567	22	AAK40160
12	801	31.0	567	22	AAI20932
13	801	31.0	567	22	AAI46176
14	801	31.0	567	22	AAI06643
15	801	31.0	567	24	AB514215
16	716	27.7	473	22	ABA50193
17	716	27.7	473	22	ABA68128
18	716	27.7	473	22	ABA51512
19	716	27.7	473	22	AAK16512
20	716	27.7	473	22	AAK42265
21	716	27.7	473	22	AAI48339
22	716	27.7	473	22	AAI08696
23	654	25.3	533	21	AAAC01342
24	565	21.8	549	24	ABK44114
25	552	21.3	2025	23	ABL05371
26	503	19.4	420	22	ABA45056
27	503	19.4	420	22	ABA55529
28	503	19.4	420	22	ABA52539
29	503	19.4	420	22	AAK03760
30	503	19.4	420	22	AAK29224
31	503	19.4	420	22	AAI35180
32	503	19.4	420	22	AAI03692
33	503	19.4	454	22	ABA42995
34	503	19.4	454	22	ABA53411
35	503	19.4	454	22	ABA23187
36	503	19.4	454	22	AAK01681
37	503	19.4	454	22	AAK27129
38	503	19.4	454	22	AAI11718
39	503	19.4	454	22	AAI33025
40	503	19.4	454	22	AAI01646
41	503	19.4	454	24	AB501685
42	502.5	19.4	899	23	ABLI2097
43	498	19.2	318	21	AAAC21875
44	481	18.6	1752	22	AA533672
45	481	18.6	1752	24	AAI38753

ALIGNMENTS

RESULT 1
AAAX26552
ID: AAAX26552 standard; DNA: 1528 BP.

XX AAAX26552;

DT 14-JUN-1999 (first entry)

XX DNA encoding chicken c-ery protein.

DE Chicken; C-11 protein; cell calcification inhibiting activity;

KW cell calcification inhibiting agent; c-ery protein; arthritis deformans;

KW ossification; spinal column ligament; ss.

XX Gallus sp.

OS Key Location/Qualifiers
FH CDS 63..1499

```

FT      /*tag= a
XX      JP11075871-A.
XX      23-MAR-1999.
XX      29-MAY-1998; 98JP-0166076.
XX      20-JUN-1997; 97US-0050297.
XX      18-JUN-1997; 97US-0878177.
XX      (CHUS ) CHUGAI PHARM CO LTD.
XX      (UYPE-) UNIV PENNSYLVANIA.
XX      WPI: 1999-257708/22.
XX      P-PSDB: AAY01521.
XX      An active protein for inhibiting cell calcification - useful for
XX      measuring the calcification of a cell, for diagnosing arthritis
XX      deformans or ossification of spinal column ligament
XX      PS      Disclosure: Page 8-9; 15pp: Japanese.
XX      CC      The present sequence encodes a chicken c-ery protein. The specification
XX      CC      also describes a chicken C-11 protein (AA01520) which has cell
XX      CC      calcification inhibiting activity and a cell calcification inhibiting
XX      CC      agent containing c-ery protein. The proteins are used for measuring the
XX      CC      calcification of a cell, for diagnosing arthritis deformans or
XX      CC      ossification of spinal column ligament.
XX      SO      Sequence 1528 BP; 464 A; 395 C; 330 G; 339 T; 0 other:

Alignment Scores:
Pred. No.:      1,95e-192      Length:      1528
Score:          2436.00      Matches:      458
Percent Similarity: 91.97%      Conservative: 0
Best Local Similarity: 91.97%      Mismatches: 0
Query Match:    94.13%      Indels:      40
DB:             20      Gaps:        2

US-09-902-772-4 (1-478) x AAX26552 (1-1528)
QY      1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB      63 ATGGCAAGCACTAATTAAAGAGCACTTATCACTGCTGAGTGAAGCAACCTCTGTTGAG 122
QY      21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB      123 TGTCCTACGAGATGCCCCACCTTGCAAGACAGAAATGACAGCTCTCTCCAGTGA 182
QY      41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPLeuSerGlnPro 60
DB      183 TATGGCAACATCAAGAGATGAGCCCGCGCTTCCAGAGAGAGCTGTTATCATCACCCC 242
QY      61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValaAsnGlySerArgAsn 80
DB      243 CCGGCCAGAGTTACATTAAAGATGAGTGAACCAACCAAGGTTAATGGGTCAAGGAAT 302
QY      81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAspVal 100
DB      303 TCACCTGATGCTGACAGCGTGGCAAAAGAGGAAATGGTTACAGATTCAAGCAATGTT 362
QY      101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProPropProAsnMetThrThr 120
DB      363 GGGATGAACTATGGAACCTCATGGAAGAGAGATATTCGCGCTCCAATATATACAAACC 422
QY      121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
DB      423 AATGAACGAAGAGATTATGTGCCAGCAGATCTACGTTATGAGACACAGACATGTAACG 482
QY      141 GlnTyrLeuGlnTyrProAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
DB      483 CAGTGGCTGGAGTGGGCAAGAGAGTAAATGTTCTTCACAGCGTGAACATCTGTTGTTTC 542

```

```

QY      161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB      543 CAGAACATGATGGGAAAGAGTTGTAAATATGACCAAAAGATGACTTCCAGAGACTCAGC 602
QY      181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB      603 CCGAGCTATTAACGAGATATCTCTGTCACACCTACACTACCAAGAGAGACTCTCTT 662
QY      201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB      663 CCACATTTGACTTATGATGATGTTGATTAAGGCTTACAAAACCTCCACGGTTAATGCAT 722
QY      221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
DB      723 GCTAGAAACACAGAGAGACGACATTTATTTTCCAAATACATCAGTTTACCAGAGCA 782
QY      241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTyr 260
DB      783 ACGCAAAAGATTAACAAAGAGCCGAGATTACCTTATAGCAAGCAGAGATCAGCGTGG 842
QY      261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
DB      843 ACGAGTACAGCCATCCCACTCACTCAAAAGCTAACCAACCATCATCTTACACAGTCCC 902
QY      281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
DB      903 AAAACAGAGACGCGCTCCATGATTGATGATCTTATATGATGATGATGATGATGATGAT 962
QY      301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuLeuLeu 320
DB      963 GCTCTTCAAAATCCAGGAGAGTGGGCAATGATGATGATGATGATGATGATGATGATGAT 1022
QY      321 SerAspSerSerAsnSerAsnCysIleThrTyrProGluGlnThrAsnGlyLysPheLysMet 340
DB      1023 TCGGACAGCTTCCAACTCACTCACTCACTGAGAGAGGACCAAAATGGGAGTTCCAAATG 1082
QY      341 ThrAspProAspGluValAlaArgArgTyrPheGlyLysSerLysProAsnMetAsn 360
DB      1083 ACAGACCTGTATGAAAGTGGCTCGGCGTTGGGAGAGAGAGAAAGCAAACTCAACATGAA 1142
QY      361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
DB      1143 TATGACAACTCAGCGCTGCACTTGCCTACTATGACAAATAATATATGACTTAAAGTT 1202
QY      380 ----- 380
DB      1203 CATGCTAAACGCTATGCTTCAAAATTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCT 1262
QY      381 HisProProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSerTyr 400
DB      1263 CACCTCCAGAAATCATCATGATGTAACAAATATACCATCAAGACCTCCCTCATGAGTCCCTAC 1322
QY      401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
DB      1322 ----- 1322
QY      421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
DB      1323 CATGCAACACCCCAAGAGATGAATTTGTAGCTTCCCATTCCTCCCTGCTTTGCCCGTAACC 1382
QY      441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPheSerProThrGlyGlyIleTyr 460
DB      1383 TCATCCAGCTTTTGTGTCCTTAATCCATCTGGAATTCACCACTGAGGACATCTAC 1442
QY      461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
DB      1443 CCCAATATACAGGCTGCGAGCTGCTCATATGCTTCCATCTTCCAGCTTACTATAC 1496

```

RESULT 2

AAX26551 standard; DNA: 1447 BP.

XX

AC AAX26551;
 XX 14-JUN-1999 (first entry)
 DE DNA encoding chicken C-11 protein.
 XX Chicken: C-11 protein; cell calcification inhibiting activity;
 KW cell calcification inhibiting agent; c-ery protein; arthritis deformans;
 KM ossification; spinal column ligament; ss.
 XX Gallus sp.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 63..1418
 FT /tag= a
 XX JP11075871-A.
 XX
 PD 23-MAR-1999.
 PD
 XX 29-MAY-1998; 98JP-0166076.
 PE
 XX 20-JUN-1997; 97US-0050297.
 PR 18-JUN-1997; 97US-0878177.
 XX
 PA (CHUS) CHUGAI PHARM CO LTD.
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 DR WPI: 1999-257708/22.
 DR P-PSDB: AAY01520.
 XX
 PT An active protein for inhibiting cell calcification - useful for
 PT measuring the calcification of a cell, for diagnosing arthritis
 PT deformans or ossification of spinal column ligament
 XX
 PS Disclosure: Page 7-8; 15pp; Japanese.
 XX
 CC The present sequence encodes a chicken C-11 protein which has cell
 CC calcification inhibiting activity. The specification also describes
 CC a cell calcification inhibiting agent containing c-ery protein
 CC (AAY01521). The proteins are used for measuring the calcification of a
 CC cell, for diagnosing arthritis deformans or ossification of spinal column
 CC ligament.
 CC
 XX
 SO Sequence 1447 bp; 440 A; 374 C; 317 G; 316 T; 0 other;
 Alignment Scores:
 Pred. No.: 4,38e-178 Length: 1447
 Score: 2262.50 Matches: 430
 Percent Similarity: 86.35% Conservative: 0
 Best Local Similarity: 86.35% Mismatches: 1
 Query Match: 87.42% Indels: 67
 DB: 20 Gaps: 3
 US-09-902-772-4 (1-478) x AAX26551 (1-1447)
 QY 1 MetAlaSerThrIleuysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 63 ATGGCAAGCACTAATTAAGAGACATATACAGTGTGAGTGAAGACCAAGTCCCTTTGAG 122
 QY 21 CysAlaTyrGlySerProHisIleuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
 DB 123 TGTGCTACGAGATGCCCCACCTTGCAGAAAGACAGAAATGACAGCCTCTTCCAGTGA 182
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
 DB 183 TATGGCAACATCAAGATGAGAGCGCGCTTCCAGACAGAGACTGTTATACACAGCCC 242
 QY 61 ProAlaArgValThrIleuLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 243 CCGGCCAGAGTTACATTAAAGATGAGTGAACCCAAACAGGTTAATGGGTCAAGCAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnVal 100

DB 303 TCACCTGATGCTCAGCGCTGCAAAAGAGAGAAATGCTTACAGCTTCAGACATTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThr 120
 DB 363 GGGATGAACATATGAGAGCTACATGAGAGAGCAATATTCGCTCCAAATATGACAACC 422
 QY 121 AsnGluArgValIleValProAlaAspProThrIleuThrPheSerThrAspHisValArg 140
 DB 423 AATGACGAAAGATTAATGTGGCCAGAGATCTACGTTATGAGACACAGACATGACGG 482
 QY 141 GlnThrLeuGluThrPalaValAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 DB 483 CAGTGGCTGAGAGTGGCAGCTCAAGAGATGCTTCCAGACGCTGAGATCTTGTGTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 543 CAGAAACATTCATGCGAAAGAGTGTGTAAATGACCAAAAGATGACTTCCAGACATCAG 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisTyrLeuArgGluThrProLeu 200
 DB 603 CCGAGCTATAACGAGATATCTCTCTCAGACCTACACTACCTCAGAGAGA----- 656
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 DB 656 ----- 656
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 DB 657 -----GGAGCCACATTTATTTTCCAAATACATCAGTTTACCAGAAAGCA 701
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 260
 DB 702 ACGCAAGAAATPAACAAGAGCCAGATTTACTTATGAGCAAGAGAGATACGCGTGG 761
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValPro 280
 DB 762 ACGAGTCAACACCCATCCACATCAACAAACCTACCATCATCTTCAACAGTGGCC 821
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 DB 822 AAACAGAGAGACCGCGCTCTCAGTTAGATCTTATGAGATTCATGAGTTCGACCCAGCAGC 881
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeu 320
 DB 882 CGTCTTGCAATTCAGGAGGTGGCAGATACAGCTATGCGAGTCTTACAGCTTCTG 941
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTyrPheGluGlyThrAsnGlyGluPheLysMet 340
 DB 942 TCGGACAGCTCCAACTCCAACTGATCATCAGTGGAGGCAAAATGGGAGTTTCAAGATG 1001
 QY 341 ThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsn 360
 DB 1002 ACAGACCCCTGATGAGTGGCTCGCGCTTGGGAGAGAGAGAAACCAAACTTAACATGAC 1061
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 DB 1062 TATGACAAACTCAGCCCGTCTGCTACTACTATGACAAATAATTTATATGCTTAAAGTT 1121
 QY 380 ----- 380
 DB 1122 CATGCTAAACGCTATGCTCAAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCT 1181
 QY 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 DB 1182 CACCTCCAGAAATCATCATGATCAAAATACCAATCAGACCTCCCTCATGATGTTCTTAC 1241
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 DB 1241 ----- 1241
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProValThr 440

Db 1242 CATGCACACCCCGAGAGATGACTTGTAGCTCCCACTCCCTGCTTCCCGTAAACC 1301
 Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLylleTyr 460
 Db 1302 TCATCCAGCTTTTGGTGGCCCTAATCCATCTGAGATTCACCACTGGAGCATCTAC 1361
 Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 Db 1362 CCCAATACAGAGTCGCCAGCTGTCATATAGCCCTTCCCATCTTGGACCTACTAC 1415
 RESULT 3
 ID AAQ50644 standard; cDNA; 2938 BP.
 AC AAQ50644;
 DT 26-MAY-1994 (first entry)
 DE Human Hum-F11-1 gene clone BM025.
 KW Chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;
 KW Ews gene; malignant melanoma; hum-f11-1;
 KW primitive peripheral neuroectodermal tumour; human chromosome 11;
 KW human chromosome 22; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 143..1501
 FT //tag= a
 FT /product= HUM-F11-1
 FT polyA_signal 2908..2913
 FT //tag= b
 PN MO9323549-A.
 XX 25-NOV-1993.
 XX 19-MAY-1993; 93WO-FR00494.
 XX 20-MAY-1992; 92FR-0006123.
 XX (CNRS) CNRS CENT NAT RECH SCT.
 PA Aurias A, Delattre O, Desmaze C, Melot T, Peter M;
 PI Plougastel B, Thomas G, Zucman J;
 DR WPI: 1993-386580/48.
 DR P-PSDB: AAR44556.
 XX
 PT New nucleic acid of EWS gene and its hybrid(s) - contg. gene
 PT sequence involved in chromosomal trans-location, also derived
 PT mRNA, probes, fusion proteins etc., for diagnosis and treatment
 PT of Ewing sarcoma and melanoma
 XX
 PS Disclosure; Fig 7; 123pp; French.
 CC The probe 11R1 was used to screen a human marrow cDNA library
 CC (Clontech cat. # HL1058). The clone BM025 was identified and
 CC sequenced. It represents the entire coding region together with
 CC 5'- and 3'-UTRs of the Hum-F11-1 gene.
 XX
 SQ Sequence 2938 BP; 847 A; 692 C; 675 G; 724 T; 0 other;
 Alignment Scores:
 Pred. No.: 5.77e-113 Length: 2938
 Score: 1482.00 Matches: 291
 Percent Similarity: 68.65% Conservative: 55
 Best Local Similarity: 57.74% Mismatches: 80
 Query Match: 57.26% Indels: 78
 DB: 14 Gaps: 9
 US-09-902-772-4 (1-478) x AAQ50644 (1-2938)

Oy 1 MetaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 143 ATGACGCGGACTATTAAAGAGGCTCTGTGGTGGTACGACGACGACGCTCTTTGAC 202
 Oy 21 CysAlaTyrGly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
 Db 203 TCACGCTAGCGAGGCGACGACCATCTCCCAAGCCGACATGATGCTGCTGGAGAGCT 262
 Oy 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59
 Db 263 GACTACGCGGCCGCCAAGATCATACCCCTCCACACGACGAGGATGATCAACAG 322
 Oy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
 Db 323 ---CCAGTGAAGGTCACAGTCACACGGAGAT---GACCAATGATGATGATCCAGG 373
 Oy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
 Db 374 GAGTCTCCGCTGAGCTCAGCGTTAGCAAAATGACAGAAAGCTGGTGGCGAGCGAGCTCC 433
 Oy 100 ValGlyMetAsnTyrGlySerTyrMetGluGluLysHis---IleProProAsnMet 118
 Db 434 AACCCCATGAACATAACAGCTATATGACGAGAGAAAGATGACCCCTCTCCCAACATG 493
 Oy 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuThrSerThrAspHis 138
 Db 494 ACCACCAACGAGAGAGATCATGTCTCCGACACCCACACACGTGTGACACAGAGACAT 553
 Oy 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeu 158
 Db 554 GTGAGGCAATGGCTGGAGTGGCCATTAAGAGATGATGCTGAGGAGATGACACATCC 613
 Oy 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
 Db 614 TTTTCCAGACATGATGGCAAGAACTGTGTAATAATGACAGAGGAGACTTCTCTCCG 673
 Oy 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThr 198
 Db 674 GCCACCAACCCCTACACAGAGAGAGTGTGTTGTCACACCTCAGATTCTCAGGAAAGT 733
 Oy 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
 Db 734 TCACCTG----- 739
 Oy 219 MethHisAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
 Db 740 -----CTGGCTTATANTACACCTCCACACC 766
 Oy 239 GluAlaThrGlnArgIleThrArgProAspLeuProTyrGluGlnAlaArgArgSer 258
 Db 767 GACCAATCCCTCAGCATGAGTGTCAAGAAGACCCCTTATATGACATCAGTCAAGAGAGA 826
 Oy 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
 Db 827 GCATGGGGCAATACATGAAATTCGTGGCTCAACAAAAGTCTCCCTGGAGGGGCAAC 886
 Oy 278 ThrValProLysThrGluAspGlnArgProGlnIleAspProTyrGlnIleLeuGlyPro 297
 Db 887 ACATATCGTAAAGATACAGAGCAAGGCGCCACACAGATCCGATATGATCCTCGGGCCG 946
 Oy 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317
 Db 947 ACCAGCAGTCGCTAGCCACACCTCGAAGCGGCGACATCCAGCTGAGCAATTCCTCTG 1006
 Oy 318 GluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluTyrThrAsnGlyGlu 337
 Db 1007 GAGCTGCTCTCCACACGCGGACGCGCAGCTGTATACCTGGAGGAGGACCAACGCGGAG 1066
 Oy 338 PheLysMetThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysPro 357
 Db 1067 TTCAAATAGACGAGACCCCGATGAGTGCGCCAGGCGCTGGGCGAGGAGAAACCAAGCCC 1126

OY 358 AsnMetAsnTyrAspLysIleuSerArgAlaLeuArgTyrTyrAspLysAsnIleMet 377
 |||||
 Db 1127 AACATGATATACGACAGCTGAGCGGCCCTCGTTATATCTATGATTAATAACATATATG 1186
 |||||
 OY 378 ThrIysValHisProProGluSerSerMetTyrTyrProSerAspLeuProTyrMet 397
 |||||
 Db 1187 ACCAAAGG----- 1195
 OY 398 SerSerTyrHisGlyLysArgTyrAlaTyrTyrLysPheAspPheHisGlyIleAlaGlnAla 417
 |||||
 Db 1196 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCCACGGCATTTGCCAGGCT 1246
 OY 418 LeuGlnPro----- 420
 |||||
 Db 1247 CTGAGAGCAATCCGACGAGTCGTCATGTACAAAGTACCTCTGCACATCTCCATACATG 1306
 OY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
 |||||
 Db 1307 CCTTCCTACACATGCCACGACAGAGGTGAATCTTGCTCCCTCCCATCCATCCCTCATG 1366
 OY 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGly 457
 |||||
 Db 1367 CCTGTCACTTCTCTCAGCTTTTGGAGCCGCAATCAATCTGACCTCCGCCACGGGG 1426
 OY 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
 |||||
 Db 1427 GGAAATACACCCCAACCCCAAGCTGCCCGGCATCTTAACACCCAGTGCCTTACACTTA 1486
 OY 475 GlyThrTyrTyr 478
 |||||
 Db 1487 GGCAGCTACTAC 1498
 RESULT 4
 ABR84139
 ID ABR84139 standard; cDNA: 2957 BP.
 AC ABR84139;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #710.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US30821.
 XX
 PR 03-OCT-2000; 2000US-237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX
 PS Claim 1; SEQ ID No 710; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA. M2 is useful for
 CC modulating GCA. M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection, and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX

SO Sequence 2957 BP; 855 A; 698 C; 680 G; 724 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	2957
Score:	1482.00	291
Percent Similarity:	68.65%	Conserved:
Best local Similarity:	57.74%	Mismatches:
Query Match:	57.26%	Indices:
DB:	24	Gaps:
		9

US-09-902-772-4 (1-478) x ABR84139 (1-2957)

OY	1	MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu	20
Db	173	ATGGACGGGACTATTAAAGAGGCTCTGCTGTGAGCAGCAGCCAGTCCCTTTGAC	232
OY	21	CysAlaTyrGly--SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer	39
Db	233	TCAAGCTACGAGGAGCGGAGCCCATCTCCCAAGGCCACATGACTGCTCGGGAGTCT	292
OY	40	GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln	59
Db	293	GACTACGGGAGGCCCAACCATCCCTCCACACAGCAGAGGTGATCAATCAG	352
OY	60	ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg	79
Db	353	--CCAGTGGAGGTCACAGCGGAGTAT-----GACCACATGATGATCCAGG	403
OY	80	AsnSerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsn	99
Db	404	GAGTCTCGGTGACATGACGCTTACGAAATGACAGCAAGCTGGTGGGAGCGAGCTCC	463
OY	100	ValGlyMetAsnTyrGlySerTyrMetGluGlnLysHis--IleProProProAsnMet	118
Db	464	ACCCCATGAATACACAGCTATATGAGCAGAGAAGATGGCCGCCCTCTCCACATG	523

```

QY 119 ThrThrasngluarGargValIleValProAlaAspProThrLeuTrpSerThrAspHis 138
D 524 ACCACCAAGAGAGAGATGATGTCGCCGACACCCACACGCTGGACACAGAGCAT 583
QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeu 158
D 584 GTGAGGCAATGGCGGAGTGGCCATTAAGAGATGACGCTTGAAGAGATCGACACATCC 643
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuGlySerMetThrLysAspAspPheGlnArg 178
D 644 TTTTTCAGACATGATGGATGGCAAGAACTGTAAATGAACCAAGAGAGACTTCCTCCGC 703
QY 179 LeuThrProSerTyraAsnAlaAspIleLeuSerHisLeuHisTyrlLeuArgGluThr 198
D 704 GCCACCAACCTCTACACACAGAGAGAGTGTGTGCACACCTCATTCCTCAGGAGAGAGT 763
QY 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
D 764 TCACGTG----- 769
QY 219 MetHisAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
D 770 -----CTGGCCTAATAATACAACTCCACACC 796
QY 239 GluAlaThrGlnArgIleThrArgProAspLeuProTyrgLuglnAlaArgArgSer 258
D 797 GACCAATCTCACAGATGAGTGCACAAAGAACCTCTATATGACTCAGTCAGAGAGAGA 856
QY 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
D 857 GCTTGGGGCAATACATGAATCTCGCCCTCAACCAAAAGTCCCTCCCTGGAGGGGCACAA 916
QY 278 ThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrgLuglnIleLeuGlyPro 297
D 917 ACCATCGTAAGATACAGAGCAAGCGCCCGACGCAATCCGATTCGATCCTGGGCCG 976
QY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317
D 977 ACCAGCAGTCGCTAGCCCAACCTCGAAGCGGCGAGATCCAGCTGTGCAATTCCTCCTG 1036
QY 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrProGluGlyThrAsnGlyGlu 337
D 1037 GACGTGCTCTCCAGACGCGCCAGCGCTGATACCTGGGGGAGGCCAACAGGGGAG 1096
QY 338 PheLysMetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysPro 357
D 1097 TTCAAAATGACGAGACCCCGATGAGTGGCCAGCGCTGGGGCGGCGGAAAAGCAAGCCC 1156
QY 358 AsnMetAsnTyraLysLeuSerArgAlaLeuArgTyrlTyraLysAsnIleMet 377
D 1157 AACATGAAATTAACACAAAGCTGACCGCGGCTCCGTTATATGATATAAACAATATATG 1216
QY 378 ThrLysValHisProGluSerSerMetTyrlLysTyrlProSerAspLeuProTyrMet 397
D 1217 ACCAAAGTG----- 1225
QY 398 SerSerTyrlHisGlyLysArgTyraLysTyrlLysPheAspPheHisGlyIleAlaGlnAla 417
D 1226 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCCACGGCATTCGCCAGGCT 1276
QY 418 LeuGlnPro----- 420
D 1277 CTCGACGCAATCCGACCGAGTGCATGTACAAAGTACCCTTCTGACATCTCCTACATG 1336
QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
D 1337 CCTTCTACATGCCACGACGAGAGGTAACCTTGTCTCCCTCCATCATCTCATCTCATC 1396
QY 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrlTyraLysSerProThrGly 457
D 1397 CCGTGAATCTTCCCTCCAGCTTCTTGGAGCGCATACAAATATCTGAGCTCCGCCACGGGG 1456

```

```

QY 458 GlyIleTyrlProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
D 1457 GGAAATCTAACCCCAACCCCAAGAGTCCCGCCGATCTTAACACCCAGTGGCTTACACTTA 1516
QY 475 GlyThrTyrlTyr 478
D 1517 GGCAGCTACTAC 1528
RESULT 5
AAH02915
ID AAH02915 standard; DNA; 2954 bp.
XX
AC AAH02915;
XX
DE 15-JUN-2001 (first entry)
XX
DE Human shear stress-response coding sequence SEQ ID NO: 83.
XX
KW Human; shear stress-response protein; vascular disease;
KW arteriosclerosis; ds.
XX
OS Homo sapiens.
XX
PN MO200125427-A1.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000MO-JP06840.
XX
PR 01-OCT-1999; 99JP-0280976.
XX
PA (KYOM ) KYOMA HAKKO KOGYO KK.
PA (NOJIT/) NOJIMA H.
PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
DR WPI: 2001-266308/27.
XX
P-PSDB: AAB90792.
XX
PT DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis -
XX
PS Claim 20; Page 462-466; 678bp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension.
XX
SO Sequence 2954 BP; 846 A; 698 C; 683 G; 727 T; 0 other;
Alignment Scores:
Pred. No.: 1.64e-111 Length: 2954
Score: 1464.50 Matches: 290
Percent Similarity: 68.45% Conservative: 55
Best Local Similarity: 57.54% Mismatches: 80
Query Match: 56.59% Indels: 79
DB: 22 Gaps: 10
US-09-902-772-4 (1-478) x AAH02915 (1-2954)
QY 1 MetAlaSerThrIleLysGluAlaLeuSerValSerGlnAspGlnSerLeuPheGln 20
D 173 ATGAGCGGACTATTAAGAGAGCTGTGCTGTGTAAGCAGACAGATCCTCTTGGAC 232
QY 21 CysAlaTyrlGly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
D 233 TCAGCGTACGAGGAGCGGACGCCATCTCCCAAGCGCAGATGATGCTCTGGGAGATCCT 292
QY 40 GluTyrlGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGln 59

```

```

Db      293 GACTACGGGCGGCCCAAGATCAACCCCTCCACACAGCAGAGTGATCAATCAG 352
      60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
      353 ---CCAGTAGAGGTCACAGTCACACCGGGAGTAT-----GACCACATGAATGATGATCAGG 403
      80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
      404 GAGTCTCGGTGGTGGAGCTGACGTTAGCAAAATGACAGACGCTGGGGCGGAGCGAGTCC 463
      100 ValGlyMetAsnTyrGlySerTyrMetGluGlnLysHis---IleProProProAsnMet 118
      464 AACCCCATGAACTCAACACAGCTATATGACAGAGAAGATGGCCCCCTCCTCCCAACATG 523
      119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTyrPseThrAspHis 138
      524 ACCACCAAGAGAGAGAGTATCTGTCCTCCGACAGCCACACTGTGGACACAGAGCAT 583
      139 ValArgGlnTrpLeuGlnTrpAlaValLysGlnTyrGlyLeuProAspValAspIleLeu 158
      584 GTGAGGCAATGGCTGGAGTGGCCATAAAGAGTACAGCTTGATGAGATGACACATCC 643
      159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
      644 TTTTTCAGAAATGATGATGGCAAGCAACTGCTAAATGAACAAGAGAGATTCCTCCGC 703
      179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLysIleTyrLeuArgGluThr 198
      704 GCCACCACTCTCTACACACGAGAGTCTGTTCACACCTCATTACCTCAGGGAAGT 763
      199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
      764 TCACTG----- 769
      219 MethAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
      770 -----CTGGCCTTATATACAACTCCACACCC 796
      239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSer 258
      797 GACCAATCTCAGTATGAGTGTCAAGAACCTTCTTATGACTCAGTCAGAGAAGA 856
      259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
      857 GCTTGGGGCAATACATGAAATCTGGCCTCAACAAGATCCTCCCTTGGAGGGGCACAA 916
      278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
      917 ACGATCAGTAAAGATACAGACCAAGGCCCCAGCCAGATCCGATACAGATCCTGGCCCG 976
      298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317
      977 ACCAGCAGTCCGCTAGCCAACTGGAAGCGGAGAGATCCAGCTGGCAATTCCTCCTG 1036
      318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyGlu 337
      1037 GAGCTGCTCCGACAGCGCCAAAGCAGCTGTATCACTGGGAGGGAGCCAAACGGGAG 1096
      338 PheLysMetThrAspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysPro 357
      1097 TTCAAAATGACGAGCCCGATGAGTGGCCAGGGCGTGGGCGAGCGGAAAAGCAAGCCC 1156
      358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
      1157 AACATGAATATGACAGAGCTAGCGGGCCCTCCGTTATCTATGATAAAACATATATG 1216
      378 ThrLysValHisProProGlnLysSerMetTyrLysTyrProSerAspLeuProTyrMet 397
      1217 ACCAAAGTG----- 1225
      398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417

```

```

Db      1226 -----CACGGCAAAAGATATGCTTACAAATTTGACTTCACAGGCATTGCCAGGCT 1276
      418 LeuGlnPro----- 420
      1277 CTGAGCCACACATCCGACCGAGTGGTCCATGTACAAGTACCTTGTACATCTCCTACATG 1336
      421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
      1337 CTTCCCAACATGCGCCACAGCAGAGAGTGACTTGTCCCTCCCACTCCATCCTCCATG 1396
      438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGly 457
      1397 CCTGTCACTTCTCCTCAGCTTCTTGTGAGCCGACACAAATACGTGAGCTCC---ACGGGG 1453
      458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
      1454 GGAATCTAACCCCAACCCCAACGTCGCCCGCATCTTAACACCCAGCTGCTTACACTTGA 1513
      475 GlyThrTyrTyr 478
      1514 GGCAGCTACTAC 1525
      Db
      Oy      475 GlyThrTyrTyr 478
      Db      1514 GGCAGCTACTAC 1525
      RESULT 6
      ID      AA050662
      AA050662 standard; DNA; 1890 BP.
      AC      AA050662;
      XX      26-MAY-1994 (first entry)
      DT      26-MAY-1994 (first entry)
      XX      Human Hum-Fil-1 gene (genomic DNA).
      DE      Human Hum-Fil-1 gene (genomic DNA).
      KW      chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;
      KW      Ews gene; malignant melanoma; hum-fil-1;
      KW      primitive peripheral neuroectodermal tumour; human chromosome 11;
      KW      human chromosome 22; ss.
      XX      OS
      XX      Homo sapiens.
      XX      Location/Qualifiers
      FH      1..39
      FT      exon
      FT      /tag= a
      FT      /number= 1
      FT      /codon_start= 22..24
      Oy      40..99
      FT      intron
      FT      /tag= b
      FT      /number= 1
      FT      /tag= c
      FT      /number= 2
      Oy      312..371
      FT      intron
      FT      /tag= d
      FT      /number= 2
      FT      /tag= e
      FT      /number= 3
      Oy      527..586
      FT      intron
      FT      /tag= f
      FT      /number= 3
      FT      /tag= g
      FT      /number= 4
      Oy      791..850
      FT      intron
      FT      /tag= h
      FT      /number= 4
      FT      /tag= i
      FT      /number= 5
      Oy      917..976
      FT      intron
      FT      /tag= j
      FT      /number= 5
      FT      /tag= k
      FT      /tag= l
      FT      exon

```

```

FT FT /number= 6
FT FT 1043..1102
FT FT **tag= 1
FT FT /number= 6
FT FT exon 1103..1162
FT FT **tag= m
FT FT /number= 7
FT FT intron 1163..1222
FT FT **tag= n
FT FT /number= 7
FT FT exon 1223..1270
FT FT **tag= o
FT FT /number= 8
FT FT intron 1271..1330
FT FT **tag= p
FT FT exon 1331..1890
FT FT /tag= q
FT FT /number= 9
XX

W09323549-A.
XX
XX
XX PD 25-NOV-1993.
XX
XX
XX PF 19-MAY-1993; 93WO-FR00494.
XX
XX PR 20-MAY-1992; 92FR-0006123.
XX
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX PA Aurias A, Delattre O, Desmaze C, Melot T, Peter M;
XX PI Plougastel B, Thomas G, Zucman J;
XX
XX WPI: 1993-386580/48.
XX
XX DR P-PSDB; AAR44556.
XX
XX
XX PT New nucleic acid of EMS gene and its hybrid(s) - contg. gene
XX PT sequence involved in chromosomal trans-location, also derived
XX PT mRNA, probes, fusion proteins etc., for diagnosis and treatment
XX PT of Ewing sarcoma and melanoma
XX
XX PS Disclosure; Fig 7 and Fig 13; 123pp; French.
XX
XX CC The intron-exon junctions of the human Ews gene and the Hum-Fl1-1
XX CC gene have been sequenced (see AAQ50646 and AAQ50662, respectively).
XX CC The different fusion products which could be formed by fusing
XX CC exons from the two genes, as happens after specific chromosomal
XX CC translocations, can be predicted (see AAQ50671-Q50678). The sequences
XX CC at fusion junctions of other observed translocations are given in
XX CC AAQ50679-Q50683). NOTE: the 1890 residue sequence AAQ50662 does not
XX CC appear in the specification; it is a combination of the Hum-Fl1-1
XX CC cDNA sequence (Fig 7, AAQ50644) and the individual intron sequences
XX CC with their intron-exon junctions (Fig 13).
XX
XX SQ Sequence 1890 BP; 457 A; 561 C; 458 G; 414 T; 0 other:

Alignment Scores:
Pred. No.: 6.12e-104 Length: 1890
Score: 1370.00 Matches: 294
Percent Similarity: 55.89% Conservative: 57
Best Local Similarity: 46.82% Mismatches: 83
Query Match: 52.94% Indels: 194
DB: 14 Gaps: 16

US-09-902-772-4 (1-478) x AAQ50662 (1-1890)

OY 5 IleYscIuAlaIeuSerValValSerGluaspGInserLeupheGluYsAlaTyrGly 24
Db 94 CTCGAGGAGGCTCTGTCGGTGGAGCGACGACAGTCCTCTTGACTCAGCGTACGGA 153
OY 25 ---SerProHisIeuAlaIylsThrGluMetThrAlaSerSerSerSerGluTyrGlyGln 43
Db 154 GCGGCAAGCCATCTCCCAAGGCGCAGACATGCTCGGGAGAGCTCGACTACGGGCGAG 213

```

OY	44	ThrsrLysMseTserProArgValProIngnInasbPTpleSerGlnProProAlaArg	63
Db	214	CCCCAAGATCAACCCCTCCACACAGCGAGGATCAATACAG---CCAGTGAAG	270
OY	64	ValThrIleLysMecGluCysAsnProAsnGlnValAsnGlySer-----	78
Db	271	GTCAACGTCAACGGGAGATAT-----GACCACATGAATGGATCCAGTAAGCTCACAG	324
OY	79	-----ArgAsnSerProAsp	83
Db	325	GCCTGTGCAGATTGGCCCTTGGCTTTGCCCTCTCCTACATTAGGAGATCTCCGGTG	384
OY	84	AspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsn	103
Db	385	GACTCAGACGCTTACGAAATGCAGCAACGCTGGTGGGGCGGAGCGGATCCAAACCCCATGAAC	444
OY	104	TyrGlySerTyrMetGlnGlnLysHis--IleProProProAsnMetThrThrAsnGln	122
Db	445	TACACACAGCTATATGGACGAGAAAGATGGCCCCCTCTCCCAACATGACACACAGAG	504
OY	123	ArgArgValIleValPro-----	128
Db	505	AGGAGAGTCATCTGCTCCCGCAGATATTCGAGAACAGGCTGCGCGCCCTCTGCT	564
OY	129	-----AlaAspProThrLeuTyrPserThrAspHisValArgGlnTyr	142
Db	555	AACAAACGCTCTCTCCTCTGCACAGACCCCACTGCTGGACACAGACATGTGACGAAATGG	624
OY	143	LeuGlnTrrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsn	162
Db	625	CTGGAGTGGCCATTAAGAGATATATGCTTGATGAGATCGACACATCTTTTCCAGAAC	684
OY	163	IleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSer	182
Db	685	ATGGATGGCAAGAACTGTGTAAATGAACAAGAGAGACTTCTCCGCGCACCAACCTCTC	744
OY	183	TyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuProHis	202
Db	745	TACAAACAGGAAAGCTCTTGTGCACACCTCAGTTACCTCAGGAAAGTAAGTGGCCGCCA	804
OY	203	LeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAlaArg	222
Db	805	AGTACC-----CAG	813
OY	223	AsnThrGlyGlyAlaThrPheIlePhe-----ProAsnThr	234
Db	814	GGCTGGGGTATTAACCTGTTATATGTTTGGCTTCAGGTTACGCTGGCCTATATATACA	873
OY	235	SerValTyrProGluAlaThrGlnArgIleThrThrArgPro-----	248
Db	874	ACCTGCCACACGACCAATCTCTCAGATGTAGTGAAGTGCACAAAGTAAGTATTTGTTCTTGG	933
OY	249	-----AspLeuProTyrGlnGln	254
Db	934	TGCACCTTAAATTTATGTACCCCTATTTGTTATTTGTCATTAGACCTCTTATATACATCA	993
OY	255	AlaArgArgSerAlaTrrPThrSerHisSerHisProThrGlnSerLys-----	270
Db	994	GTCAGAAAGAGAGCATGGGCAATACATACATGAAATTCGTGCTCAACAAAAGTAAGTAATG	1053
OY	271	-----AlaThrGlnPro-----	274
Db	1054	TTTATATAGTCTTTGGAGGCTCAGCTGCTTTCTTTCTCTCTGTCACAGGCTCTCCCTT	1113
OY	275	---SerSerSerThrValProLysThrGlnAspGlnArgProGln-----	288
Db	1114	GGAGGGGCAACAAGATCATGAATAATACAGAGCAACGGCCCAAGCTAGCTACGCCCA	1173
OY	289	-----LeuAspProTyrGln	293
Db	1174	GGATATGTATCTCTCTTTGAGCAAAATTTCTTTTATTTATCTTATGATCCGTATCAG	1233

Db 234 CCCCCGAGTACTCTGTACAGTACCCCTCAGACCTCCGATACATGGGCTCTATAC 175
 QY 422 AAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
 Db 174 GCCACCCACAGAAATGATGACTTTGTGGCGCCCAACCTCCAGCCCTCCCGTGACATCT 115
 QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 461
 Db 114 TCCAGTTTTTTTGGCGCCCAAAACCATCTGATTCACCAACTGGGGGTATATACCC 55
 QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrTyr 478
 Db 54 AACCTAGGCTGCCACCAAGCCATATGCTTCTCATCTGGGCACTTACTAC 4
 RESULT 8
 ABA66003/C
 ID ABA66003 standard; DNA; 567 BP.
 XX
 AC ABA66003;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #14308.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 4; SEQ ID NO 14308; 639bp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;

US-09-902-772-4 (1-478) x ABA66003 (1-567)
 QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuGlnLeuLeuSer 321
 Db 534 CTGACCTCATTAAGGAGAGTGGCCAGATCCAGCTTTGGAGTTCTCTCTGGACTCTCTGCG 475
 QY 332 AspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyGlnPheLysMetThr 341
 Db 474 GACAGCTCAACTCCAGCTGATCATCCCTGGGAAGGACCAACGCGGAGTTCAAGTACAG 415
 QY 342 AspProAspGlnValAlaAlaArgTrpGlnGlyLysSerLysProAsnMetAsnTyr 361
 Db 414 GATCCCGAGGAGTGGCCGCGCTGGGGAAGAGGGAAGCAAAACCAATGAAATCAC 355
 QY 362 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 381
 Db 354 GATAGCTCAGCCGCGCCCTCGTTACTACTGATGCAAGAACATCATATCAACAGGTC 298
 QY 382 ProProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401
 Db 297 -----CAT 295
 QY 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 Db 294 GGGAGCGCTAGCGCTACAGTTCAGTTCACAGGGATGCGCCAGGCTCCAGCCAC 235
 QY 421 -----His 421
 Db 234 CCCCCGAGTACTCTGTACAGTACCCCTCAGACCTCCGATACATGGGCTCTCATAC 175
 QY 422 AAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
 Db 174 GCCACCCACAGAAATGATGACTTTGTGGCGCCCAACCTCCAGCCCTCCCGTGACATCT 115
 QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 461
 Db 114 TCCAGTTTTTTTGGCGCCCAAAACCATCTGAAATTCACCACTGGGGGTATATACCC 55
 QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrTyr 478
 Db 54 AACCTAGGCTGCCACCAAGCCATATGCTTCTCATCTGGGCACTTACTAC 4
 RESULT 9
 ABA33090/C
 ID ABA33090 standard; DNA; 567 BP.
 XX
 AC ABA33090;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Probe #11556 for gene expression analysis in human heart cell sample.
 XX
 KW Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 4; SEQ ID No 11556; 530pp; English.
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.78e-57 Length: 567
 Score: 801.00 Matches: 151
 Percent Similarity: 78.17% Conservative: 3
 Best Local Similarity: 76.65% Mismatches: 3
 Query Match: 30.95% Indels: 40
 DB: 22 Gaps: 2
 US-09-902-772-4 (1-478) x ABA3090 (1-567)
 QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 321
 DB 534 CTGACCTCATTTAGGAGTGGCCAGATCCAGCTTGGCGAGTCTCTGAGCTCTGCG 475
 QY 322 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThr 341
 DB 474 GACAGCTCCACCTCAGCTGCATCAGCTGGAGGACCAACGAGGAGTCAAGTACG 415
 QY 342 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTrp 361
 DB 414 GATCCCGAGGAGGAGGCGGCGTGGGAGAGGAGGAGCAACCAACATGAACTAC 355
 QY 362 AspLysLeuSerArgAlaLeuArgTrpTrpLysAspLysAsnIleMetThrLysValHis 381
 DB 354 GATAAGCTCAGCCCGCCCTCCGTTACTACTATGACAAAGACATCATGACCAAGTTC 298
 QY 382 ProProGluSerSerMetLysTrpLysTrpSerAspLeuProTyrMetSerSerTrpHis 401
 DB 297 -----CAT 295
 QY 402 GlyLysArgTrpAlaTrpLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
 DB 294 GGGAGGCGCTACGGCTTACAGTTGACACTTCCACGAGGAGGCCCGCTCAGCCAC 235
 QY 421 -----His 421
 DB 234 CCCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGGTACATGGGCTCTATCAC 175
 QY 422 AlaHisProGlnLysMetAsnPheValAlaLabronHisProProAlaLeuProValThrSer 441
 DB 174 GCCACCCACAGAGATGATTTGTGTGGCCGCCCTCCAGCCCTCCCGGTGACATCT 115
 QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 461
 DB 114 TCCAGTTTTTTTGGCCCCCAAAACCATATGGAATTACCAACCTGGGGGATATATACCC 55
 QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
 ||||||| :|||

DB 54 AACACTAGGCTCCCGCCAGCCAGCATATGCTTCTATCTGCGACTTACTAC 4
 RESULT 10
 ID AAK14426 standard; DNA; 567 BP.
 AC AAK14426;
 DT 05-NOV-2001 (first entry)
 XX Human brain expressed single exon probe SEQ ID NO: 14417.
 DE Human: brain expressed exon: gene expression analysis; probe:
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 OS Homo sapiens.
 XX W0200157275-A2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-0500667.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483446/52.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 14417; 650pp + Sequence Listing; English.
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.78e-57 Length: 567
 Score: 801.00 Matches: 151
 Percent Similarity: 78.17% Conservative: 3
 Best Local Similarity: 76.65% Mismatches: 3
 Query Match: 30.95% Indels: 40
 DB: 22 Gaps: 2
 US-09-902-772-4 (1-478) x AAK14426 (1-567)
 QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 321
 DB 534 CTGACCTCATTTAGGAGTGGCCAGATCCAGCTTGGCGAGTCTCTGAGCTCTGCG 475
 QY 322 AspSerSerAsnSerAsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThr 341
 DB 474 GACAGCTCCACCTCAGCTGCATCAGCTGGAGGACCAACGAGGAGTCAAGTACG 415
 QY 342 AspProAspGluValAlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTrp 361
 ||||||| :|||

```

Db 414 GATCCGACGAGGTGGCCGGCGCTGGGAGACGGGAGAACCAACCACTGACTAC 355
Qy 362 AsplysleuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 381
Db 354 GATAGCTCAGCGCGCCCTCGTACTACTATGCAAGACATCATGACCAAGTTC--- 298
Qy 362 ProProgluSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHis 401
Db 297 -----CAT 295
Qy 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
Db 294 GGGAGAGCGTACGCTACAGTTCACAGTTCACAGGGAGTCCAGGCCCTCCAGCCAC 235
Qy 421 -----His 421
Db 234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTATAC 175
Qy 422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
Db 174 GCCCACCACAGAAATGACTTGTGGCGCCGCCACCTCCAGCCCTCCGCTGACATCT 115
Qy 442 SerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGlyGlyIleTyrPro 461
Db 114 TCCAGTCTTTTGTGCTGCCCAACCCATGCTGGAATTCACCACTGGGGGTATATACCC 55
Qy 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db 54 AACCTAGGCTCCCGACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 4
RESULT 11
AAK40160/c
ID AAK40160 standard; DNA; 567 BP.
AC AAK40160;
XX
XX 06-NOV-2001 (first entry)
DE Human bone marrow expressed single exon probe SEQ ID NO: 14717.
XX
XX Human: bone marrow expressed exon: gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00066.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 14717; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX

```

```

CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX SQ Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.78e-57 Length: 567
XX Score: 801.00 Matches: 151
XX Percent Similarity: 78.17% Conservative: 3
XX Best Local Similarity: 76.65% Mismatches: 3
XX Query Match: 30.95% Indels: 40
XX DB: 22 Gaps: 2
XX
XX US-09-902-772-4 (1-478) x AAK40160 (1-567)
XX
Qy 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheLeuLeuGlnLeuSer 321
Db 534 CTGACCTCATTAAGCAGTGGCCAGATCCAGCTTGGCAGTTCCTCGAGCTCCTGCG 475
Qy 322 AspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyLysPheLysMetThr 341
Db 474 GACAGCTCACTCCAGCTGACATCACCCTGGGAGGACCAACGGGAGTTCAAGATGACG 415
Qy 342 AspProAspGluValAlaArgArgTyrGlyLysLysSerLysProAsnMetAsnTyr 361
Db 414 GATCCCGACGAGGTGGCCGGCGCTGGGAGAGCGGAGACCAACCAACATGAACTAC 355
Qy 362 AsplysleuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 381
Db 354 GATAGCTCAGCGCGCCCTCGTACTACTATGCAAGACATCATGACCAAGTTC--- 298
Qy 382 ProProgluSerMetTyrLysTyrProSerAspLeuProTyrMetSerTyrHis 401
Db 297 -----CAT 295
Qy 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
Db 294 GGGAGAGCGTACGCTACAGTTCAGTTCACCGGAGTCCAGGCCCTCCAGCCAC 235
Qy 421 -----His 421
Db 234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTATAC 175
Qy 422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
Db 174 GCCCACCACAGAAATGACTTGTGGCGCCGCCACCTCCAGCCCTCCGCTGACATCT 115
Qy 442 SerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGlyGlyIleTyrPro 461
Db 114 TCCAGTCTTTTGTGCTGCCCAACCCATGCTGGAATTCACCACTGGGGGTATATACCC 55
Qy 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db 54 AACCTAGGCTCCCGACCAAGCATATGCTTCTCATCTGGGCACTTACTAC 4
RESULT 12
AAI20932/c
ID AAI20932 standard; DNA; 567 BP.
AC AAI20932;
XX
XX 12-OCT-2001 (first entry)
XX
XX Probe #10865 for gene expression analysis in human cervical cell sample.
XX
XX Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX

```


PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID No 10865; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.78e-57 Length: 567
Score: 801.00 Matches: 151
Percent Similarity: 78.17% Conservative: 3
Best Local Similarity: 76.65% Mismatches: 3
Query Match: 30.95% Indels: 40
DB: 22 Gaps: 2
US-09-902-772-4 (1-478) x AA120932 (1-567)
QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuSer 321
DB 534 CTGACCTCATTAAGCAGATGGCCACATCCAGCTTGGCAGTCCCTCGAGCTCTGCG 475
QY 322 AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMetThr 341
DB 474 GACAGCTCCCAACTCCAGCTGCATCAGCTGGAGAGCCACCAAGGGAGTTCAAGATGACG 415
QY 342 AspProAspGluValAlaArgArgTrpGluArgLysSerLysProAsnMetAsnTyr 361
DB 414 GATCCCGACGAGGGGCGGCGCTGGGAGAGCGGAGAACCAACCAATGAACTATC 355
QY 362 AspLysLeuSerArgAlaLeuArgTrpTyrTrpLysAsnIleMetThrLysValHis 381
DB 354 GATNAGCTACGCGCGCGCTCGTTACTACTATCACAAGAACATCATGACCAAGTGC--- 298
QY 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401
DB 297 -----CAT 295
QY 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
DB 294 GGGAAAGCGCTACGCTTCAAGATTGACTCCAGGGAGTCCGCGCTCCAGCCCGAC 235
QY 421 -----His 421
DB 234 CCCCCGAGTATCTGTACAAAGTACCCCTCAGACCTCCGCTACATGGGCTCTATACAC 175

QY 422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValTrpSer 441
DB 174 GCCACCCACAGAAATGATTTGTGGCCGCCACCTCCAGCCCTCCGTCGATCT 115
QY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyrPro 461
DB 114 TCCAGTTTTTTTGTGTCGCCCAACCCATCTGGAATTACGACAGTGGGGATATATACCC 55
QY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
DB 54 AACACTAGGCTCCGCCACGACCATATGCTTCTCATCTGGGACACTTACTAC 4
RESULT 13
AA146176/c
ID AA146176 standard; DNA; 567 BP.
XX
XX AA146176;
AC
XX 17-OCT-2001 (first entry)
DT
XX
XX Probe #14862 used to measure gene expression in human placenta sample.
DE
XX Probe; microarray: human; placenta; antenatal diagnosis;
KW genetic disorder: ss.
XX
XX Homo sapiens.
OS
XX WO200157272-A2.
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00663.
PE
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 14862; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.78e-57 Length: 567
Score: 801.00 Matches: 151
Percent Similarity: 78.17% Conservative: 3
Best Local Similarity: 76.65% Mismatches: 3
Query Match: 30.95% Indels: 40
DB: 22 Gaps: 2
US-09-902-772-4 (1-478) x AA146176 (1-567)
QY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuLeuSer 321
DB 534 CTGACCTCATTAAGCAGATGGCCACATCCAGCTTGGCAGTCCCTCGAGCTCTGCG 475

Db 534 CTGACCTCATTAGGAGCAGTGGCCAGATCCAGCTTTGGGAGCTTCCTCCGAGCTCCTGTCG 475
 Oy 322 AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMetThr 341
 Db 474 GACAGCTCACAATCCAGCTGCATCACCCTGGAGAGGACCAAGGGAGTTCAAGATGACG 415
 Oy 342 AspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 361
 Db 414 GATCCCGACGAGAGTGGCCCGCGCTGGGGAGAGCGGAGCAACCAACCATGATGACATAC 355
 Oy 362 AspLysLeuSerArgAlaLeuAlaIleArgTrpGlyTyrAspLysAsnIleMetThrLysValHis 381
 Db 354 GATAGCTCAGCCGGCCCTCCGTTACTACTATGACAAAGATCATCATGACCAAGATC--- 298
 Oy 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401
 Db 297 -----CAT 295
 Oy 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
 Db 294 GGGAGAGCGCTACGGCTTACAGATTCGATTCACAGGGATCCGCGCCCTCCAGCCCTCAC 235
 Oy 421 -----His 421
 Db 234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGCTCCTATAC 175
 Oy 422 AlaHisProGlnLysMetAsnPheValAlaIleProHisProProAlaLeuProValThrSer 441
 Db 174 GCCACCCACAGAGAAGATGAATCTTGTGGCGCCACCCCTCCAGCCCTCCGCTGACATCT 115
 Oy 442 SerSerPhePheAlaIleAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyrPro 461
 Db 114 TCCAGTTTTTTTGTGCCCCAACCCATCTGGAATTCACCACTGGGGGTATATACCCC 55
 Oy 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
 Db 54 AACACTAGCTCCCGCCACGCAATATGCTTCTCATCTGGGCACTTACTAC 4
 RESULT 14
 AA106643 standard; DNA; 567 BP.
 AC AA106643;
 XX 09-OCT-2001 (first entry)
 DE Probe #6634 used to measure gene expression in human breast sample.
 XX Probe: human; breast disease; breast cancer; development disorder; ss.
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX Homo sapiens.
 OS WO200157270-A2.
 XX 09-AUG-2001.
 PD 29-JAN-2001; 2001WO-US00661.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.

PR Novel single exon nucleic acid probe used to measuring gene expression
 PR in a human breast -
 XX
 PS Claim 25; SEQ ID NO 6634; 322pp; English.
 CC The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 567 BP; 114 A; 124 C; 200 G; 129 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.78e-57 Length: 567
 Score: 801.00 Matches: 151
 Percent Similarity: 78.17% Conservative: 3
 Best Local Similarity: 76.65% Mismatches: 3
 Query Match: 30.95% Indels: 40
 DB: 22 Gaps: 2
 US-09-902-772-4 (1-478) x AA106643 (1-567)
 Oy 302. leuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuSer 321
 Db 534 CTGACCTCATTAGGAGCAGTGGCCAGATCCAGCTTTGGGAGTTCCTCCGAGCTCCTGTCG 475
 Oy 322 AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMetThr 341
 Db 474 GACAGCTCACAATCCAGCTGCATCACCCTGGAGAGGACCAACGCGGAGTTCAATGATGACG 415
 Oy 342 AspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsnTyr 361
 Db 414 GATCCCGACGAGAGTGGCCCGCGCTGGGGAGAGCGGAGCAACCAACCATGATGACATAC 355
 Oy 362 AspLysLeuSerArgAlaLeuAlaIleArgTrpGlyTyrAspLysAsnIleMetThrLysValHis 381
 Db 354 GATAGCTCAGCCGGCCCTCCGTTACTACTATGACAAAGATCATCATGACCAAGATC--- 298
 Oy 382 ProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401
 Db 297 -----CAT 295
 Oy 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
 Db 294 GGGAGAGCGCTACGGCTTACAGATTCGATTCACAGGGATCCGCGCCCTCCAGCCCTCAC 235
 Oy 421 -----His 421
 Db 234 CCCCCGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGCTCCTATAC 175
 Oy 422 AlaHisProGlnLysMetAsnPheValAlaIleProHisProProAlaLeuProValThrSer 441
 Db 174 GCCACCCACAGAGAAGATGAATCTTGTGGCCCCAACCCCTCCAGCCCTCCGCTGACATCT 115
 Oy 442 SerSerPhePheAlaIleAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyrPro 461
 Db 114 TCCAGTTTTTTTGTGCCCCAACCCATCTGGAATTCACCACTGGGGGTATATACCCC 55
 Oy 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
 Db 54 AACACTAGCTCCCGCCACGCAATATGCTTCTCATCTGGGCACTTACTAC 4
 RESULT 15
 ABSI4215/c

THIS PAGE BLANK (USPTO)

12 299.5

— ()

123 TGTGCTACGATCGCCCACTTGCAAGACAGAAATGACAGCCTCCTCTTCCAGTGAA 182

QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
 |||||
 DB 183 TATGGGCAACATCAAAAGATGAGCCCGCGCTCCCGCAGAGAGACTGCTATCAACAGCCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 |||||
 DB 243 CCGGCGAAGATTACATTAAAGATGAGTGTAAACCAACCAAGTTAAATGGGCAAGGAAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 |||||
 DB 303 TCACCTGATGACTGAGCGGTGGCAAAAGAGGAAATGGTTAGAGTTCCAGACAAATTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
 |||||
 DB 363 GGGATGAACATGTGGAAGCTACATGGAAGAGAAAGATATCCGCCCTCCAAATATGACAAAC 422
 QY 121 AsnGluArgValIleValProAlaAspProThrLeuTyrPheSerThrAspHisValArg 140
 |||||
 DB 423 AATGAACGAGAGATTATTGTGCCAGCAGATCTCAGTTATGAGACACAGACCATGTACGG 482
 QY 141 GlnTyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
 |||||
 DB 483 CAGTGGCTGGAGTGGCGCAGTGAAGAGATGTGCTTCACAGCTGGACATCTTGTGTTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 |||||
 DB 543 CAGAACATTTGATGGAAAGAGTGTGTAAATGACCAAAAGATGACTTCAGAGACTCAGC 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
 |||||
 DB 603 CCGAGCTATAACGAGATATCTCTCTGACACCTACACCTACACCTGAGAGACACTCTCTT 662
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 |||||
 DB 663 CCACATTTGACTTGATGATGATGTGATGATGAGCCCTTACAAACTCTCCAGCTTAATGAT 722
 QY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 |||||
 DB 723 GCTGGAACACAGAGAGAGCCACTTTATTTCCTCAATACATCATGTTACCCAGAAAGCA 782
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyr 260
 |||||
 DB 783 ACGGAAGAATTAACAACAGAGCCAGATTTACCTTATGAGCAGAGAGATGACAGCGTGG 842
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 |||||
 DB 843 ACGAGTCACAGCCATCCACATCACTCAAAAGCTAACACCATCATCTTCAACAGTGGCC 902
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 |||||
 DB 903 AAAACAGAGACCCAGCGCTCTCACTAGATCTTATGAGATTTTGGACGACAGCAGCAGC 962
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGluLeuLeu 320
 |||||
 DB 963 CGCTTTCCAATCCAGGAGAGGAGGAGATACAGTATGAGCAAGTTCCTCACTGAGCTTCTG 1022
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTyrGlnGlyThrAsnGlyGluPheLysMet 340
 |||||
 DB 1023 TCGGACAGCTCCAACTCCAACTGATCATCTCGGAGAGGACCAAAATGGGAGTTAAAGATG 1082
 QY 341 ThrAspProAspGluValAlaArgArgTyrPheGlyLysArgLysSerLysProAsnMetAsn 360
 |||||
 DB 1083 ACAGACCTGATGAAGTGGCTCGGCGTGGGAGAGAGAGAAACCAAACTAACATGAAAC 1142
 QY 361 TyrAspLysLysSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 |||||
 DB 1143 TATGACAACTCAGCCGCTGCACTTCTGCTACTACTATGACAAAATATATGACTAAAGTT 1202
 QY 380 ----- 380
 DB 1203 CATGGTAAAGCTATGCTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCT 1262

QY 381 HisProProGluSerSerMetLysTyrTyrProSerAspLeuProTyrMetSerSerTyr 400
 |||||
 DB 1263 CACCTCCAGATCATCTCATGTACAAATACCATCAGACCTCCCTACATGATGTTCTAC 1322
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 |||||
 DB 1322 ----- 1322
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 |||||
 DB 1323 CATGACACCCCGCAAGATGACATCTGTACTGCTCCCATCCCGCTTTCGCCGTAC 1382
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrTrpAsnSerProThrGlyLysTyr 460
 |||||
 DB 1383 TCATCCAGCTTTTGTGCGCCCTTAATCCATATGGAATTCACCAACGAGGACATCTAC 1442
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 |||||
 DB 1443 CCAATATACAGGCTGCCAGCTGCTCATATGACTTCCCATCTTGGACCTACTAC 1496
 RESULT 2
 US-08-878-177-1
 ; Sequence 1, Application US/08878177
 ; Patent No. 6294354
 ; GENERAL INFORMATION:
 ; APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
 ; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
 ; TITLE OF INVENTION: the Proteins
 ; FILE REFERENCE: chugai selyaku kabushiki kaisha 5001
 ; CURRENT APPLICATION NUMBER: US/08/878,177
 ; CURRENT FILING DATE: 1997-06-18
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1447
 ; TYPE: DNA
 ; ORGANISM: C-11 gene, c-erg gene w/ deletion, chicken DNA
 ; US-08-878-177-1
 Alignment Scores:
 Pred. No.: 1, 28e-219 Length: 1447
 Score: 2262.50 Matches: 430
 Percent Similarity: 86.35% Conserved: 0
 Best Local Similarity: 86.35% Mismatches: 1
 Query Match: 87.42% Indels: 67
 DB: 4 Gaps: 3
 US-09-902-772-4 (1-478) x US-08-878-177-1 (1-1447)
 QY 1 MetalaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGlu 20
 |||||
 DB 63 ATGCAGACACTATTAAAGAAAGCTTATACAGTGTGATGTAAGACCAAGTCTTGTGAG 122
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 |||||
 DB 123 TGTCCCTACAGATGCCGCCCACTTGCACAAAGACAAATATGACAGCCTCTCTCCAGTGA 182
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
 |||||
 DB 183 TATGGCAACATTAAGATGAGCCCGCGGCTTCCCGACGAGGAGCTGTTATACACAGCCC 242
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 |||||
 DB 243 CCGGCGAAGATTACATTAAAGATGAGTGTAAACCAACCAAGTTAAATGGGCAAGGAAAT 302
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 |||||
 DB 303 TCACCTGATGACTGAGCGGTGGCAAAAGAGGAAATGGTTAGAGTTCCAGACAAATTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
 |||||
 DB 363 GGGATGAACATGTGGAAGCTACATGGAAGAGAAAGATATCCGCCCTCCAAATATGACAAAC 422

QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
 Db 423 AATGAAGAGAGATTATGTCGCCAGAGATCTCAGTTATGAGACAGACCATGTCAGG 482
 QY 141 GluThrLeuGluThrPalalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
 Db 483 CAGTGGCTGGAGTGGAGTGAAGAGATGATGCTTCACAGCGTGGACATCTGTGTCTC 542
 QY 161 GluAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 543 CAGACACTGTATGGGAAAGATTGTGTAATGACCAAGATGACTTCCAGAGACTCAGC 602
 QY 181 ProSerThrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
 Db 603 CCGAGCTATACGAGATATCTCTCTCAGACCTACACTACCTCCTCAGAGACAGA- 656
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 Db 656 ----- 656
 QY 221 AlaArgAsnThrGlyLysIleThrPheIlePheProAsnThrSerValTyrProGluAla 240
 Db 657 -----GAGCCACTTATATTTTCCAAATCATCAGTTTACCCAGAGCA 701
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTyr 260
 Db 702 ACGCAAGAAATTAACAAGAGCCGATTTACCTTATGACGAGAGAGATCAGCTGG 761
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 Db 762 ACGAGTCACAGCATCCCTCAGTCAAAAGCTACCAACCATCATCTTCAACAGTCC 821
 QY 281 LysThrGluAspGluArgProGlnLeuAspProTyrGlnIleLeuGlnProThrSerSer 300
 Db 822 AAACACAGAGACGCGCTCTCAGTTAGTCTTATCAGATCTTGAGCCAGCAGCAGC 881
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuGluLeuLeu 320
 Db 882 CGTCTTCCAATCCAGGAGATGGGACATACAGTATGCGAGTTCCTACTGGAGCTTCTG 941
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTyrGluGlyThrAsnGlyLysPheLysMet 340
 Db 942 TCGGACAGCTCCAACTCAATGTCATCCTGGAGGAGCAAAATGGAGTTCAGATG 1001
 QY 341 ThrAspProAspGluValAlaArgArgTyrGlyLysSerLysProAsnMetAsn 360
 Db 1002 ACAGACCTGATGAAGTGGCTCGGCTGGGAGAGAGAAACCAACCTTAACATGAC 1061
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 Db 1062 TATGACAAACTCAGCCGCTGCATCTGCTACTACTATGACAAAATATTATGACTAAAGTT 1121
 QY 380 ----- 380
 Db 1122 CATGGTAAACGCTATGCTCAAAATTTGATTTCCACGGAATCGCTCAGCCCTCAGCT 1181
 QY 381 HisProProGluSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 Db 1182 CACCTCTCAGATATCATGATCAATTAACCAATCAACCTCCCTTCATGAGATGCTCTAC 1241
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 Db 1241 ----- 1241
 QY 421 HisAlaHisProGluLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 Db 1242 CAGTCAACACCCAGAGATGAACTTGTAGTCCCATCCCTCGCTTGGCCGTAAC 1301
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrArgLysIleTyr 460
 Db 1302 TCATCCAGCTTTTGTGCTGCCCTTAATCCATACGGAATTCACCAACTGGAGGCACTTAC 1361
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478

Db 1362 CCCAATACAGGCTGCCAGCTGCATATATGCTTCCACTCTTGACACTACTAC 1415
 RESULT 3
 US-08-343-443B-3
 : Sequence 3, Application US/0834343B
 : Patent No. 5968734
 GENERAL INFORMATION:
 : APPLICANT: Aurias, Alain
 : APPLICANT: Delattre, Olivier
 : APPLICANT: Desmaze, Chantal
 : APPLICANT: Melot, Thomas
 : APPLICANT: Peter, Martine
 : APPLICANT: Plooungastel, Beatrice
 : APPLICANT: Thomas, Gilles
 : APPLICANT: Zucman, Jessica
 : TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
 : TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
 : TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
 : TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
 : TITLE OF INVENTION: TRANSLATIONS
 : NUMBER OF SEQUENCES: 129
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Weiser & Associates
 : STREET: 230 South Fifteenth Street
 : CITY: Philadelphia
 : STATE: PA
 : COUNTRY: USA
 : ZIP: 19102
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: AEDIT 1.0 DOS text editor
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/343,443B
 : FILING DATE: 18-NOV-1994
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/FR93/00494
 : FILING DATE: 19-MAY-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: FR 92/06123
 : FILING DATE: 20-MAY-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Weiser, Gerard J.
 : REGISTRATION NUMBER: 19,763
 : REFERENCE/DOCKET NUMBER: 989,6121P
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 215-875-8383
 : TELEFAX: 215-875-8394
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2938 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 143..1498
 : US-08-343-443B-3
 Alignment Scores:
 Pred. No.: 4.88e-140 Length: 2938
 Score: 1482.00 Matches: 291
 Percent Similarity: 68.65% Conservative: 55
 Best Local Similarity: 57.74% Mismatches: 80
 Query Match: 57.26% Indels: 78
 DB: 2 Gaps: 9
 US-09-902-772-4 (1-478) x US-08-343-443B-3 (1-2938)
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20


```

Db      996 CTCT-----TACTGGCCTGTCCCAAGCCACCGCGCTGC 1031
Oy      456 -----ThrlgyllyllyrProAsnThrArgLeuPro----- 466
Db      1032 CGCGCCCGCCGCTGCGCTTACCCCAAGCCGCGCTGCAGCCCTCCGCGGCGCTTTGG 1091
Oy      467 -AlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1092 CGCGGTGGCGCGCGCTTGCACCTTGGGGGCTCATTTAT 1128

RESULT 6
US-09-344-579-1
: Sequence 1, Application US/09344579
: Patent No. 6054316
: GENERAL INFORMATION:
: APPLICANT: Brenda F. Baker
: APPLICANT: Lex M. Cowser
: TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
: FILE REFERENCE: RTS-0063
: CURRENT APPLICATION NUMBER: US/09/344,579
: CURRENT FILING DATE: 1999-06-25
: NUMBER OF SEQ ID NOS: 47
: SEQ ID NO 1
: LENGTH: 2268
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (291)..(1700)
US-09-344-579-1

Alignment Scores:
Pred. No.: 6,71e-35      Length: 2268
Score: 442.50           Matches: 133
Percent Similarity: 39.50%      Conserves: 40
Best Local Similarity: 30.37%    Mismatches: 108
Query Match: 17.10%           Indels: 158
Gaps: 16

US-09-902-772-4 (1-478) x US-09-344-579-1 (1-2268)
Oy      122 GUAARGAYVALLIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db      570 CAGGGGCGCGGCGGCGATTCGCAAGAACCCCTGCGCTGGAGTGAAGCAACAGGTATGCCAG 629
Oy      142 TTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db      630 TGCGTCTCTGGGCCACCAATGATGATGCTGTGTGAACGTGAATGTCAGAGGTTCC-- 686
Oy      162 AsnIleAspGlyLysGluLeuGlyLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      687 GGCATGAATGGCCAGATGCTGTGTAACTTGGCAAGACGCTTTCGAGCTGCGACCT 746
Oy      182 SerTyrAsnAlaAspIleLeuLeuSerHisLeu----- 192
Db      747 GACTTGTGGTGACATCTCTGGAAACATCTGGAAACATGATCAAGAAAGAACAGAA 806
Oy      193 -----HisTyrLeuArg 196
Db      807 AAGACAGACATCAATATGAAGAAATTCACACCTCCTCCGTTCTCTCATTTGATTAAC 866
Oy      197 GluThr----- 198
Db      867 AGCAATACATTAGTTTGGACAGACAGCGCCCTATGATGACAGACACAGATTTAC 926
Oy      199 -----ProLeu-ProHisLeuThrSerAspAspVa 208
Db      927 CCCAAAGCGCGCTCTGTGACAGCATGTGCCGCTCCACACCCAGCTACTGAGCTCT 986
Oy      208 LAspLysAlaLeuGlnAsnSerProArgLeu----- 218
Db      987 GAGCAGAGATTTCAGATGTTCCCAAGTTCGCGCTCAGTCCGCTCAGCGCTACCTACTGC 1046

```

```

Oy      219 -MetHisAlaArgAsnThrGlyAlaThrPheIle-----Ph 231
Db      1047 TCTGTACGTACAGGACTTCCAGCGACGACACTTGATTTGCTCACCACAAATTCGGGACT 1106
Oy      231 eProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgProAspLeuPr 251
Db      1107 CCCAAAGACCGAGACT--CCCTGAGAAACGGTCCGCGACAGCTTCAGAGAGCTCAGACTCC 1163
Oy      251 oTyrGluGln-----AlaArgArgSerAlaTrp----- 260
Db      1164 CTCCTCCAGTCCGGAACAGCAGTCTGCTGTGCTGATGCAAGGGTCTCTCCTTC 1223
Oy      261 -----ThSerHisSerHisProThrGlnSe 269
Db      1224 GAGAGCTTCGAAGATGACGTGACGACAGTCTCTGCTCCTCAATGAACCAACCATCTCTTC 1283
Oy      269 rLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgProGlnLe 289
Db      1284 AAGATTACA-----TCCAAGAGAGAGAGTACCCAGTG-CAGCAAGCAAAACAGTTAT 1336
Oy      289 uAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyI 309
Db      1337 ACTGCAGCTGTGCTGGCGGCTTCACA-----GGAAGTGAGCC 1375
Oy      309 nIleGlnLeuTrpLysPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnCySII 329
Db      1376 TATTCAGCTGTGAGAGTTTCCCTGAGAGCTCTATCAGACAAATCCGCGCATTCATCAT 1435
Oy      329 eThrTyrPheGluThrAsnGlyLysPheLysMetThrAspProAspGluValAlaArgAr 349
Db      1436 CAGCTGAGCTGGAACAGGATGGAGTTTAAAGCTGCGGACCCCATAGGTGGCGCCCGC 1495
Oy      349 gTyrGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuAr 369
Db      1496 GTGGGGAAGAGGAAATTAAGCCCAAGATGAATACAGAGAGCTGAGCGGCGCTTACG 1555
Oy      369 gTyrTyrTyrAspLysAsnIleMetThrLysValHisProGluSerMetTyrLys 389
Db      1556 CTACTATTACGACAAAGAACATCATCATCAAGAGAGC----- 1589
Oy      389 sTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrAlaTyrLysph 409
Db      1590 -----TCGGGGAAGCGCTACGTACCCCTT 1615
Oy      409 e-----AspPheHisGlyIleAlaGln 416
Db      1616 CGTGTGGACCTCCAGACTTGTCTGGGTTTCACGCGCCGAGAGAACTGCACGCGCATCTG-- 1673
Oy      416 nAlaLeuGlnProHisAla-----HisProGlnLysMetAsnph 429
Db      1674 -GGCGTCCAGCGCCGACACGAGAGCTGAGTCCGCGGACACACCTGAGCGG----- 1724
Oy      429 eValAlaPro-----HisProAlaLeuPro 438
Db      1725 ---GCCCAAGGCTGTGAGCTGAGTGAAGGACCAATCTGACCAAGCTGCT 1772

RESULT 7
US-08-306-691B-43
: Sequence 43, Application US/08306691B
: Patent No. 5734039
: GENERAL INFORMATION:
: APPLICANT: Calabretta, Bruno
: APPLICANT: Skorski, Tomasz
: TITLE OF INVENTION: ANTISENSE
: TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
: CORRESPONDENCE ADDRESS: 55
: ADDRESSEE: Seidel, Gonda, Lavorqua & Monaco, P.C.
: STREET: Two Penn Center, Suite 1800
: CITY: Philadelphia
: STATE: Pennsylvania
: COUNTRY: U.S.A.

```

```

; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-306-691B-43

```

```

Alignment Scores:
Pred. No.: 6,09e-35 Length: 1604
Score: 440.50 Matches: 115
Percent Similarity: 44.21% Conservative: 53
Best Local Similarity: 30.26% Mismatches: 102
Query Match: 17.02% Indels: 111
Gaps: 8

```

US-09-902-772-4 (1-478) x US-08-306-691B-43 (1-1604)

```

QY 119 ThThrAsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138
   ||| :|||:||||| :|||: ||||| :|||: |||
Db 447 ACTAAGAACACGACGACGAGGCGGAGAAAGACCCCGCGGACGACGAAACCCAT 506

QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValaAspIleLeu 158
   ||||| |||||:||||| |||||: ||| ||||| |||||
Db 507 GTTCGGAGCTGGGTGATGCTGCTGAATGAATTCAGCCTGAAGGTGAGACTTCACG 566

QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
   ||| :|||:||||| |||||: ||| ||||| |||||
Db 567 AAGTTC---TGATGAATGAGACGACGCTGCGCCCTGGGTAAAGACTCTTCGACG 623

QY 179 LeuThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThr 198
   ||| ||| :|||:||||| |||||: ||||| |||||
Db 624 CTGGCCCAAGACTTGTGGGACATCTTATGGACATCTAGAGATCCTGCAGAAAGAG 683

QY 199 ProLeuPro-HisLeuThrSerAspAspValaAspLysAlaLeuGlnAsn----- 214
   :|||: |||||: |||||: |||||: |||||: |||||
Db 684 GATGTAAACCATATCAAGTTAATGAGTCACACCCATCCATCCATCCATCCATAC 743

QY 215 -----SerProArgLeuMetHisAlaIleAsnTh 224
   :|||: |||||: |||||: |||||: |||||: |||||
Db 744 TCGGATTACTTCATAGCTATGATTTAGACATGCCACAGTGTCTCCACATGGAGTTC 803

QY 224 rGlyGlyAlaThrPheIle----- 230
   :|||: |||||: |||||: |||||: |||||: |||||
Db 804 TCAGAGCCC-AGCTTATTCACAGAGTCCTATCAGACGCTCATCCATCCATCCGAGA 862

QY 231 -----PheProAsnThrSerVal----- 236
   :|||: |||||: |||||: |||||: |||||: |||||
Db 863 GCTCTCTCTCCCTCAAGTATGAGATGATACCCCTGCTCATCTCTCCGAGACCCCTCCA 922

QY 237 -----TyrProGlnAlaThrGlnArgGlyLeuThrThrArgProAs 249
   ||| ||| :|||: |||||: |||||: |||||: |||||

```

```

Db 923 GACAGACACCTTCAGAAATGACTATTGCTATCAAAAGAAAGTGTACACCCAGACAA 982
QY 249 PleuProTyrGlnGlnAlaIleArg----- 257
   :|||: |||||: |||||: |||||: |||||: |||||
Db 983 CATGTGCATGGGAGGACGACGATGCTGTAAACTCGGGGCGGACACTCTTTGAACCAT 1042

QY 258 -SerAlaTrpThrSerHisSerHisProThrSerSerArgLeuAlaAsnProGlySerGlyGln 277
   :|||: |||||: |||||: |||||: |||||: |||||
Db 1043 AGAGAGCTACGATAGTGTATGATCGCCCTACCCAGTCTCTGAGAGCAGCATCTTTCAA 1102

QY 277 rThrValProLys-----ThrGlnAsp----- 284
   :|||: |||||: |||||: |||||: |||||: |||||
Db 1103 CAGCGTCACAGCTGTCTCTCTCTATAGACACTTCAGACTCAGAGAGCATATCCGGTCCCT 1162

QY 285 -----GlnArgProGlnLeuAs 290
   :|||: |||||: |||||: |||||: |||||: |||||
Db 1163 GCCCAACACACACCCAGGACGACCTTCAAGACTATGTCGGGACGCTGCTGACCTCAA 1222

QY 290 pProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGln 310
   :|||: |||||: |||||: |||||: |||||: |||||
Db 1223 TAAGGACACCTGTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1282

QY 310 eGlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerSerAsnSerAsnGlyLeu 330
   ||||| ||||| ||||| ||||| ||||| |||||
Db 1283 CCAAGTATGACGATTTCTTCTGGAATTAACACGATTAATCCTGTCAGTCTTTATCAG 1342

QY 330 rTrpGluGlyThrAsnGlyLysGluPheLysMetThrAspProAspGluValaIleArgGtr 350
   :||| ||| :|||: |||||: |||||: |||||: |||||
Db 1343 CTGACAGGAGATGCTGCTGGAATTCAACTTCTGACCCAGATGAGTGTGCGCAGAGATG 1402

QY 350 pGlyLeuArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyr 370
   ||||| ||||| ||||| ||||| ||||| |||||
Db 1403 GGGAAAGAGGAAACAAACCTTAAGATGAATTTAGAAACATGAGCCGTGCTTACCTTA 1462

QY 370 rTyrTyrAspLysAsnIleMetThrLysValHisProProGlnSerSerMetTyrLysTyr 390
   ||||| ||||| ||||| ||||| ||||| |||||
Db 1463 CTATTACGACAAAACATATCATCCACAGACAA----- 1493

QY 390 rProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrAlaTyrLysPhe 409
   :|||: |||||: |||||: |||||: |||||: |||||
Db 1494 -----GGGGGGAACGCTAGCTGATCGCGTTT 1520

RESULT 8
PCT-US93-06251-9
; Sequence 9, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366

```


QY 422 -----AlaHisProGlnIlyMetAspPheValAlaProHisProProAlaL 437
 Db 487 GGTGGCTGGGGGTGACAGTCCCGCAGAGT-----GCCCGCCAG 525
 QY 437 euProValThrSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProthG 457
 Db 526 TGGCGTGGGTGTAGCCACTTCGCTCCCTCC-----TCACGCCCTCCG 573
 QY 457 TgTllyleTyrProAsnThrArgLeuProAlaAlaHisMetProSer---HisLeu 474
 Db 574 A-GGTGCTGTCCCGCCACCG-AGGACCCCGCTCACACCGCTGCTCTTCATCTT 627

RESULT 11
 US-08-343-443B-5
 : Sequence 5, Application US/08343443B
 : Patent No. 5968734
 : GENERAL INFORMATION:
 : APPLICANT: Aurias, Alain
 : APPLICANT: Delattre, Olivier
 : APPLICANT: Desmaze, Chantal
 : APPLICANT: Melot, Thomas
 : APPLICANT: Peter, Martine
 : APPLICANT: Ploougastel, Beatrice
 : APPLICANT: Thomas, Gilles
 : APPLICANT: Zucman, Jessica
 : TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
 : TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
 : TITLE OF INVENTION: TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
 : TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
 : NUMBER OF SEQUENCES: 129
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Weiser & Associates
 : STREET: 230 South Fifteenth Street
 : CITY: Philadelphia
 : STATE: PA
 : COUNTRY: USA
 : ZIP: 19102
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: AEDIT 1.0 DOS text editor
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US//08/343,443B
 : FILING DATE: 18-NOV-1994
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/FR93/00494
 : FILING DATE: 19-MAY-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: FR 92/06123
 : FILING DATE: 20-MAY-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Weiser, Gerard J.
 : REGISTRATION NUMBER: 19,763
 : REFERENCE/DOCKET NUMBER: 989, 6121P
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 215-875-8383
 : TELEFAX: 215-875-8394
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 328 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 1..327
 : US-08-343-443B-5

Alignment Scores:

Pred. No.: 5,57e-22 Length: 328
 Score: 301.50 Matches: 57
 Percent Similarity: 79.31% Conservative: 12
 Best Local Similarity: 65.52% Mismatches: 17
 Query Match: 11.65% Indels: 1
 DB: 2 Gaps: 1

US-09-902-772-4 (1-478) x US-08-343-443B-5 (1-328)

QY 252 TyrGlnGlnAlaArgSerAlaTrpThrSerHisSerHisProthGlnSerIlyAla 271
 Db 67 TATGACTCAGTCAGAAAGAGAGAGCTTGGGCAATACATGATTTGGCCTTAACAAAAGT 126
 QY 272 ThrGln---ProSerSerThrValProIlyThrGlnAspGlnArgProGlnLeuAsp 290
 Db 127 CCTCCCTTGGAGGGGCGCAACAAACATGATGAAATACAGACAGCAACGGCCACAGAT 186
 QY 291 ProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyGlnIle 310
 Db 187 CCGTATCAGATCGGGGCGGACGACGAGTGCCTACCAACCTTGGAAGGGGCGAGATC 246
 QY 311 GlnLeuTrpGlnPheLeuLeuGlnLeuSerAspSerSerAsnSerAsnCysIleThr 330
 Db 247 CAGCTGTGCAATTCCTCTCTGAGACTGCTCTCCACAGCCGCAACGCCAGCTGATACAC 306
 QY 331 TrpGlnGlyThrAsnGlyIle 337
 Db 307 TGGAGGGGAGCAACGGGGAG 327

RESULT 12
 US-08-875-944B-1
 : Sequence 1, Application US/08875944B
 : Patent No. 6096542
 : GENERAL INFORMATION:
 : APPLICANT: FUJINAGA, Kei
 : APPLICANT: YOSHIDA, Koichi
 : APPLICANT: HIGASHINO, Fumihiko
 : TITLE OF INVENTION: CANCER CONTROL
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BROWDY AND NEIMARK, P. L. L. C.
 : STREET: 624 Ninth Street N.W., Ste. 300
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : ZIP: 20001
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/875,944B
 : FILING DATE: 07-AUG-1997
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: JP 07-020173
 : FILING DATE: 08-FEB-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/JP96/00016
 : FILING DATE: 09-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BROWDY, Roger L.
 : REGISTRATION NUMBER: 25,618
 : REFERENCE/DOCKET NUMBER: 737-3528
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 628-5197
 : TELEFAX: (202) 737-3528
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2064 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single

US-09-902-112-4 (1-4/8) X US-08-8/5-944B-1 (1-2064)

US-09-116-049-3

```
; Sequence 3, Application US/09116049A
; Patent No. 6248351
```

GENERAL INFORMATION:

APPLICANT: Hung, Men-Chie

;; TITLE OF INVENTION: HUMAN PEAS IS A TUMOR
FILE REFERENCE: TMCO-583

FILE REFERENCE: UISC:362
CURRENT APPLICATION NUMBER: MS/09/116 0492

CURRENT FILING DATE: 1998-07-15

NUMBER OF SEO ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 206

```

; TYPE: DNA

```

; ORGANISM: H
ne-09-116-049-3

03-03-010-00-00

Alignment Scores:

Pred. No.:	1.77e-20
------------	----------

Score: 299.50

Percent Similarity: 34.84%

Best Local Similarity: 26.778

Query Match: 11.57%

DB: 4

TELEFAX: (512) 474-7577
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2410 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 us-08-780-835B-1

Alignment Scores:

Pred. No.: 2.42e-18 Length: 2410
 Score: 279.50 Matches: 108
 Percent Similarity: 35.22% Conservative: 41
 Best Local Similarity: 25.53% Mismatches: 115
 Query Match: 10.80% Indels: 160
 Gaps: 15

US-09-902-772-4 (1-478) x US-08-780-835B-1 (1-2410)

QY 45 SerLysMetSerProAlaGValProGlnGlnAspThrLeuSerGlnProProAlaArgVal 64
 Db 634 TCAGAGAGGAAAGCCAGAGTCCCTCC-----GCACAGACCCCGCTGCTCCT 678
 QY 65 ThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAsp-- 83
 Db 679 GCACAGAGAGACCCACCACTCCCTACCACTGAGAGAGACAGTCCCTTACTCCAGACAAA 738
 QY 84 -----AspCysSerValAlaLys 89
 Db 739 TCCGCATCAAGTCCCGCTCCGCTCCCTGACAGAGTCCCGCTGACGCTTTTCCA 798
 QY 90 GLyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrglySerTyrglyMetGlu 109
 Db 799 GGGCAGAAACAGACAGACAGCTCC-----TGAGAG 828
 QY 110 GlnLysHisIleProProAsnMetThrThrAsnGlnArgValIleValProAla 129
 Db 829 CCTCCAGCTCTCCCACTCCACCTGAGCCAGGCTGCTGAGACACAGCTCCCTCT 888
 QY 130 AspProThrLeuThrSerThrAspHisValArgIleThrLeuGlnIleValValLysGlu 149
 Db 889 TCCAGCAGCCCGTGACA----- 906
 QY 150 TyrglyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLysGluLeuCys 169
 Db 907 -----TGT 909
 QY 170 LysMetThrLysAspAspPheGlnArgLeuThrProSerTyrrAsnAlaAspIleLeuLeu 189
 Db 910 -----GCCACTCTCTCA----- 921
 QY 190 SerHisLeuHisTyrlLeuArgGluThrProLeuProHisLeuThrSerAspAspValAsp 209
 Db 922 -----CATCTCTCAGAGAG--GGGGCCGGGAACCTCTCCAGCCCCCTATCAAC 969
 QY 210 LysAlaLeuGlnAsnSerProArgLeuMetHisAlaArgAsnThrGlyAlaThrPhe 229
 Db 970 ACCAAGTGTGGAGCCCTGCCACCTACCCCGGAGAACTTCAACAGAGAGTACATG 1029
 QY 230 IlePheProAsnThrSerValTyrrProGlnAlaThrGlnArgIleThrThrArgProAsp 249
 Db 1030 -----ACCCCTGTAGCAAGCGCTGCCAGC----- 1056
 QY 250 LeuProTyrglnGlnAlaArgArgSerAla-----Trr 260
 Db 1057 -----CCGCTT---CAAGCCAGGGTGGGTCACTGGGACACAGTACCCAGGGGGGTGG 1110
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 Db 1111 TGATCAACAAGAGAGCGACAGACTTCCCTACAGACTCAGATGTCCTGATGTGCATCAA 1170
 QY 281 LysThrGlnAspGlnArgProGlnLeuAspProTyrglnIle----- 294

Db 1171 TGTACTCCACCAGAGGGCTTCTGTGACCCCTCTCCAGGTGATGATGGATTATG 1230
 QY 295 -----Leu 295
 Db 1231 GCTATGAAAATCCCTTCGACCATTCACAGATGATGTCTGATTTGCTCCATAAAATTG 1290
 QY 296 GlyProThrSerSerArg-----LeuAlaAsnProGly----- 306
 Db 1291 AAGGAGACATCAACAGAGAGAGATGAGCTTCCGGAG-GGGCCACCTTACCACCGC 1349
 QY 307 SerGlyGlnIleGlnIleThrPheLeuLeuLeuLeuLeuSerAspSerSerAsnSer 326
 Db 1350 CGGGGTCCCTTACACAGTGTGACAGTTCGTGTGCTCCCTGCTGATGACCCCAAAATGCT 1409
 QY 327 AsnCysIleThrTrpGlnGlyThrAsnGlyGluPheLysMetThrAspProAspGluVal 346
 Db 1410 CATTTCATTCGTTGGACAGCGCGGGAATGAGTTAACTTAATTGAACCTGAAGAGTT 1469
 QY 347 AlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsnTyrrAspLysLeuSerArg 366
 Db 1470 GCCAGGCTCTGGGTATCCAGAAAGACCGGCCACCATGATATATGACAAGCTGAGCCGC 1529
 QY 367 AlaLeuArgTyrrTyrrAspLysAsnIleMetThrLysValHisProProGlnSerSer 386
 Db 1530 TCGCTGGATACATATTATGAAAGGCATCATCGAAGGTG----- 1571
 QY 387 MetTyrrLysTyrrProSerAspLeuProTyrrMetSerSerTyrrHisGlyLysArgTyrrAla 406
 Db 1572 -----GCTGGCAAGCTTACGTCG 1589
 QY 407 TyrrLysPhe 409
 Db 1590 TACAAATTT 1598

Search completed: July 28, 2003, 07:41:48
 Job time : 79.3165 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 06:34:34 ; Search time 301.516 Seconds

(without alignments)

3270.525 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKKALSVSDSLFE.....IYPTKRLPAHMSHLGTY 478

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09902772/runat.23072003_093658_14962/app_query.fasta_1.1294
-DB=Published.Applications.NA -QFMT=fastap -SUFFIX=inp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09902772@cgn.1.1.82@runat.23072003_093658_14962
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONLOG -DEV_TIMEOUT=120 -WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published.Applications.NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	2436	94.1	1528	11 US-09-902-772-3

2	2262.5	87.4	1447	11	US-09-902-772-1	Sequence 1, Appl
3	2180.5	84.3	3166	15	US-10-205-823-98	Sequence 98, Appl
4	801	31.0	567	10	US-09-864-761-18410	Sequence 18410, A
5	716	27.7	473	10	US-09-864-761-20472	Sequence 20472, A
6	638	24.7	467	12	US-09-918-995-23356	Sequence 23356, A
7	565	21.8	549	10	US-09-923-779-54	Sequence 54, Appl
8	503	19.4	420	10	US-09-864-761-3705	Sequence 3705, Ap
9	503	19.4	454	10	US-09-864-761-1653	Sequence 1653, Ap
10	481	18.6	1752	10	US-09-850-799-1	Sequence 1, Appl
11	444	17.2	2188	11	US-09-920-3004-1716	Sequence 1716, Ap
12	444	17.2	2188	11	US-10-033-528-1716	Sequence 1716, Ap
13	444	17.2	2268	11	US-09-920-3004-1693	Sequence 1693, Ap
14	444	17.2	2268	15	US-10-033-528-1693	Sequence 1693, Ap
15	443.5	17.1	3692	15	US-10-106-698-824	Sequence 824, App
16	442.5	17.1	2269	11	US-09-954-531-955	Sequence 955, App
17	430.5	16.6	1884	11	US-09-925-300-420	Sequence 420, App
18	382	14.8	348	10	US-09-864-761-29364	Sequence 29364, A
19	382	14.8	477	10	US-09-864-761-5035	Sequence 5035, Ap
20	382	14.8	519	10	US-09-864-761-12798	Sequence 12798, A
21	377	14.6	2532	15	US-10-037-270-869	Sequence 869, App
22	374	14.5	553	10	US-09-864-761-7427	Sequence 7427, Ap
23	371	14.3	225	10	US-09-864-761-21761	Sequence 21761, A
24	311	12.0	473	12	US-09-918-995-27904	Sequence 27904, A
25	305	11.8	3178	15	US-10-108-605-124	Sequence 124, App
26	305	11.8	3178	15	US-10-108-605-128	Sequence 128, App
27	303	11.7	416	12	US-09-918-995-8472	Sequence 8472, Ap
28	299.5	11.6	2064	12	US-09-884-363-3	Sequence 128, App
29	299.5	11.6	2064	15	US-10-167-485-1	Sequence 3, Appl
30	299.5	11.6	2333	11	US-09-920-3004-1788	Sequence 1, Appl
31	299.5	11.6	2333	11	US-09-880-107-3316	Sequence 1788, Ap
32	299.5	11.6	2333	15	US-10-033-528-1788	Sequence 3316, Ap
33	297.5	11.5	1894	10	US-09-126-945B-1	Sequence 1788, Ap
34	297.5	11.5	1905	10	US-09-866-356-2	Sequence 2, Appl
35	297.5	11.5	1942	15	US-10-157-031-83	Sequence 83, Appl
36	297	11.5	4106	15	US-10-084-817-187	Sequence 187, App
37	280.5	10.8	1087	15	US-10-102-806-215	Sequence 215, App
38	280.5	10.8	2952	10	US-09-925-297-257	Sequence 257, App
39	279.5	10.8	2410	12	US-09-884-363-1	Sequence 1, Appl
40	279.5	10.8	165	10	US-09-864-761-23294	Sequence 23294, A
41	278.5	10.8	1894	10	US-09-841-963A-1	Sequence 1, Appl
42	276	10.7	300	15	US-10-177-063-15	Sequence 15, Appl
43	267	10.3	2256	15	US-10-098-841-317	Sequence 317, App
44	261	10.1	300	15	US-10-177-063-16	Sequence 16, Appl
45	244.5	9.4	1426	10	US-09-925-297-309	Sequence 309, App

ALIGNMENTS

RESULT 1
US-09-902-772-3
Sequence 3, Application US/09902772
Patent No. US20020164739A1
GENERAL INFORMATION:
APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
TITLE OF INVENTION: the Proteins
FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
CURRENT APPLICATION NUMBER: US/09/902,772
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US/08/878,177
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 3
LENGTH: 1528
TYPE: DNA
ORGANISM: c-ery gene, chicken DNA
US-09-902-772-3
Alignment Scores:
Pred. No.: 4,24e-254
Score: 2436.00
Percent Similarity: 91.97%
Length: 1528
Matches: 458
Conservative: 0

Best Local Similarity: 91.97% Mismatches: 0
 Query Match: 94.13% Indels: 40
 DB: 11 Gaps: 2

US-09-902-772-4 (1-478) x US-09-902-772-3 (1-1528)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValIleSerGluAspGlnSerLeuPheGlu 20
 DB 63 ATGGCAAGCACTATTAAAGAGCATTTACAGTGGTGAAGACACAGCTCTGTTGAG 122
 QY 21 CysAlaIleTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 123 TGTGCTACGAGATGCCCCACCTTGCAAGACAGAAATGACAGCTCTCTCCAGTGA 182
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60
 DB 183 TATGGCAAAACATCAAAAGATGAGCCCGGGTCCAGACAGAGCTGTTATCACAGCCC 242
 QY 61 ProAlaArgValIleThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 243 CCGGCCAGAGTTACATTAAAGTAGAGTGAACCCAAACAGGTTAATGGTCAAGGAAT 302
 QY 81 SerProAspSerCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 303 TCACCTATGATGACGTCACGTCGCAAGAGAGGAAATGGTTACAGCTCAGACAAATGTT 362
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThr 120
 DB 363 GGGATGAACATATGGAACTCATGAGAGAGAGCATTTCCGGCTCCAAATATGACAAACC 422
 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuThrIleuThrAspHisValArg 140
 DB 423 AATGAACGAAGAGTATTGGCCAGCAGATCTACGTTATGAGACACAGCATGTACGG 482
 QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAlaIleLeuPhe 160
 DB 483 CAGTGGCTGAGTGGAGAGTGAAGAGATATGTTCTTCACAGCAGGACATCTTGTTC 542
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 543 CAGAACATGTATGGGAAAGAGTGTGTAAATGACCAAGATGACTTCCAGAGCTCAGC 602
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
 DB 603 CCGAGCTATACCCAGATATCCCTGTCACACTACACCTCAACCTCAAGAGAGACTCCCTT 662
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 DB 663 CCACATTTGACTTCAGATGATGTTGATAGAGCCCTTACAAACCTCCACGCTTATGCA 722
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 DB 723 GCTAGAAACACAGAGAGGACCACTTTATTTCCAAATCATCAGTTTACCAGAGCA 782
 QY 241 ThrGlnArgGlyLeuThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrp 260
 DB 783 ACCCAAGATTAACAACAGGAGGAGATTTACTTATGAGCAAGCAGAGATAGCGTGG 842
 QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 DB 843 ACAGATCACAGCCATCCCACTCACTCAAAAGCTAACCAACATCATCTTCAACAGTGGCC 902
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 DB 903 AAAACAGAAAGACAGCGTCTCTCAGTTAGATCTTATCAGATTTCTGGACCGACAGAGC 962
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeuLeu 320
 DB 963 CGCTTGTCAATCCAGGAGTGGGAGATACAGCTATGAGCATGCTTCCACTACGAGCTTCTG 1022
 QY 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGlnGlyThrAsnGlyLysIleLysMet 340
 DB 1023 TCGGACAGCTCCAACTCCAACTGATCCTGGGAGGGACAAATATGGGAGTTCAGATG 1082

QY 341 ThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMetAsn 360
 DB 1083 ACAGACCCCTGATGAAGTGGCTCGCGCTTGGGAGAGAGGAAAGCAAACTTAACATGAAC 1142
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 DB 1143 TATGACAACTCAGCCCTGACCTTCCCTACTACTATGACAAAATATATATGACTAAAGTT 1202
 QY 380 ----- 380
 DB 1203 CATGCTAAACGCTATGCTTCAAAATTTGATTTCCACGGAATCGCTACGCCCTCCAGCT 1262
 QY 381 HisProGlnLysSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 DB 1263 CACCCCTCAGATCATCATGATGACAAATATACCATCAGACCTCCCTCATGATGATCTTAC 1322
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisIleAlaGlnAlaLeuGlnPro 420
 DB 1322 ----- 1322
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 DB 1323 CATGACACCCCGAAGAGATGAACTTTGTAGCTCCCATGCCCTGCTTGGCCGTAC 1382
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIleTyr 460
 DB 1383 TCATCCAGCTTTTGGTGGCCCTTAATCATCATGAGAAATCACCACTGAGAGCATCTAC 1442
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 DB 1443 CCCAATACAGAGCTGCCAGCTGCATATGCTTCCATCTTGGCAGCTTACTAC 1496

RESULT 2

US-09-902-772-1
 : Sequence 1, Application US/09902772
 : Patent No. US20020164739A1
 : GENERAL INFORMATION:
 : APPLICANT: Chugai Sanyaku Kabushiki Kaisha, 5001, Iwamoto et
 : TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
 : TITLE OF INVENTION: the Proteins
 : FILE REFERENCE: chugai sanyaku kabushiki kaisha 5001
 : CURRENT APPLICATION NUMBER: US/09/902,772
 : PRIOR FILING DATE: 2001-07-12
 : PRIOR APPLICATION NUMBER: US/08/878,177
 : NUMBER OF SEQ ID NOS: 7
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 1447
 : TYPE: DNA
 : ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
 : US-09-902-772-1

Alignment Scores:

Pred. No.: 2,53e-235 Length: 1447
 Score: 2262.50 Matches: 430
 Percent Similarity: 86.35% Conservative: 0
 Best Local Similarity: 87.42% Mismatches: 1
 Query Match: 87.42% Indels: 67
 DB: 11 Gaps: 3

US-09-902-772-4 (1-478) x US-09-902-772-1 (1-1447)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValIleSerGluAspGlnSerLeuPheGlu 20
 DB 63 ATGGCAAGCACTATTAAAGAGCATTTATCAGTGTAGTGAACACAGCTCTGTTGAG 122
 QY 21 CysAlaIleTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 123 TGTGCTACGAGATGCCCCACCTTGCAAGAGAGAAATGACAGCTCTCTTCCAGTGA 182
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60

D	b	341	GCCTACGGAGCGCCACACCTGGCTTAAGACAGAGATGACCGCTCTCTCCGACGACATAT	400
Q	y	42	GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnProPro	61
D	b	401	GGACAGACTTCCCAAGATGAGCCACGCGCTCCTCCAGAGGATGGCTGCTCAACCCCA	460
Q	y	62	AlaArgValThrIleLysMetGluCysAnpProAsnGlnValAsnGlySerTrpAsnSer	81
D	b	461	GCCAGGGTCACCATCAATAATGGAAATGTAACCTTACCAGGTGAATGGCTCAAGAACTT	520
Q	y	82	ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly	101
D	b	521	CCTGATGATGACAGTGTGGCCAAAGCGGGAAGATGGTGGGAGCCACAGACCGTTGGG	580
Q	y	102	MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrTrpAsn	121
D	b	581	ATGAACACTCGGACGCTACATGAGAGGAGAGCAATGACCAACCCCAACATGACCAAGAC	640
Q	y	122	GluArgValAlaIleValProLysAspProThrLeuTyrSerThrAspHisValArgGln	141
D	b	641	GAGCCCAAGATTATGTGTGTCGACAGATCTCTACGCTATGTAGACTACAGACCATGTGGGCGAG	700
Q	y	142	TyrLeuGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln	161
D	b	701	TGGCTGGAGTGGGCGGTGAAGAAATATGGCTTCACAGCGTCAACATCTGTATTTCAG	760
Q	y	162	AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro	181
D	b	761	AACATCGATGGGAAGAACTGTGCAGATGACCAAGACAGACTTCCAAGAGCTCAACCCG	820
Q	y	182	SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro	201
D	b	821	AGCTACAAACGCGAGATCTCTCTCCATCTCCACTACTACTAGAGAGACTCTCTTCCA	880
Q	y	202	HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla	221
D	b	881	CATTGACTTACAGATGATGTGATTAACCTTACCAAACTCTCCACGGTTAATGCATGCT	940
Q	y	222	ArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluLalaThr	241
D	b	941	AGAAACACA-----	949
Q	y	242	GlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTyrThr	261
D	b	950	-----GATTATCCATATGAGCCGCCCCAGAGATCAGACCTCGAGAC	988
Q	y	262	SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro	280
D	b	989	GGTCAAGCGCCACCCACCCACCCAGTCGAAAGCTGCTCAACCATCTCTCCACACTGGCC	1048
Q	y	281	LysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer	300
D	b	1049	AAAATGTAAGACCAAGCTCTCAGTTACATCTTATCGATTTCTTGGACCAACAAGTACG	1108
Q	y	301	ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGluLeuLeu	320
D	b	1109	CGCCTTGCAATTCAGAGCATGTGGCCAGATTCAGCTTGGCATGTTCTCTGGAGCTCTCG	1168
Q	y	321	SerIlePheSerAsnSerAsnSerAsnCysIleThrTrrpGluGlyTrrpAsnGlyGluPheLysMet	340
D	b	1169	TCGGACACCTCCAACTCCAGCTGCATCCCTGGGAAGGACACCAACGGGAGTCCAAGATG	1228
Q	y	341	ThrAspProAspGluValAlaArgArgTrrpGlyGluArgLysSerLysProAsnMetAsn	360
D	b	1229	ACGATATCCCGACAGAGTGTGCCCGGCGCTGGGGAGAGCGGAAGACCAACCATGTGAC	1288
Q	y	361	TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal	380
D	b	1289	TACGATTAAGCTACAGCGCGGCTCGGTACTACTATGACAAAGAACATCATGACCAAGTGC	1348
Q	y	381	HisProProGlnSerSerMetLysTyrTyrProSerAspLeuProTyrMetSerSerTyr	400

```
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18410
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000163.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
; OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 9.00e-99
; OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00
US-09-864-761-18410

Alignment Scores:
Pred. No.: 1.8e-77 Length: 567
Score: 801.00 Matches: 151
Percent Similarity: 78.17% Conservative: 3
Best Local Similarity: 76.65% Mismatches: 3
Query Match: 30.95% Indels: 40
DB: 10 Gaps: 2

US-09-902-772-4 (1-478) x US-09-864-761-18410 (1-567)

OY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeuSer 321
|||
|||
|||
DB 534 CTGACCTATTGAGTCAGTGGCCAGATCCAGCTTGGCGAGTCCCTGACCTCTGCG 475
322 AspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGlnPheLysMetThr 341
474 GACACCTCCCACTCCAGTCGATCATCCTGGGAGGACACCAACGGGAGATTCAAGTCAGC 415
OY 332 AspProAspGluValAlaIleArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyr 361
414 GATCCCGAGAGAGTGGCCGCGGCGGAGAGCGAGAGACCAACCAACATGAACTAC 355
OY 362 AspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHis 381
364 GATAAGCTACGCCGCGCTCCGTTACTACTATGACAGAAACATCATGACCAAGGTC--- 298
OY 382 ProProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSerTyrHis 401
297 -----CAT 295
OY 402 GlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro--- 420
294 GGAAGAGCGCTACGCTCAAGTTCATCCACGCGGATGCGCCAGGCCCTCCAGGCCAC 235
OY 421 -----His 421
DB 234 CCCCCGAGATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGCGGCTCTATAC 175
422 AlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSer 441
DB 174 GCCCAGCCAGAGATCACTTTGTGGCGGCCGCCACCTCCAGGCCCTCCGCTGACACT 115
OY 442 SerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyrPro 461
DB 114 TCAACTTTTGTGGCGGCCCAAAACCATGGAATTCACCAACTGGGGGTATATACCC 55
OY 462 AsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyIleTyrTyr 478
DB 54 AACACTAGGCTCCACAGCCGATATGCTTCTCATCTGCGGCACTTACTAC 4
RESULT 5
```

```
US-09-864-761-20472/c
; Sequence 20472, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmics-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20472
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP00021.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 6.00e-88
; OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00
US-09-864-761-20472

Alignment Scores:
Pred. No.: 2.27e-68 Length: 473
Score: 716.00 Matches: 136
```


Alignment Scores:	
pred. No.:	5, 36e-60
Score:	638.00
Percent Similarity:	97.66%
Best Local Similarity:	93.75%
Length:	467
Matches:	120
Conservative:	5
Mismatches:	3

344 AspidinValAlaIargargTIpclgysguArgysSerLysProAsnmetAsnnyrAspLys 363


```

Oy 364 leuSERATGAlaAlaLeuAArgTyrTyrTyrAAspLysAsnLLeuethrLysValHisProPro 383
Db 62 CTCACCGCGCCCTCCGTACTACTATGACAGAGATCATCTGACCAAGGTC----- 112
Oy 384 GluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLys 403
Db 113 -----CATGGAG 121
Oy 404 ArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro----- 420
Db 122 CGCTACGCGCTACAGATTCACCTCCACAGGATGCGCCACGCCCTCCAGCCACCCCCG 181
Oy 421 -----HisAlaHis 423
Db 182 GAGTCATCTCTGTACAGATGACCCCTCAGACCTCCGTACATGGGCTCTATCAGGCCAC 241
Oy 424 ProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSerSerSer 443
Db 242 CCACAGAGATATACCTTGTGGCGCCCCACCCCTCCAGCCCTCCCGGAGATCTTCCAGT 301
Oy 444 PhePheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyLysIleTyrProAsnThr 463
Db 302 TTTTGTGCGCCCAACCATCTGGAATTACCAACATGGGGGTATATATACCCCTCAGT 361
Oy 464 ArgLeuProAlaAlaHisMetProSerHisLeu-GlyThrTyrTyr 478
Db 362 ANGCTCCCAACGACCATATGCTTTTCATCTTGGGACACTTACTAC 407

RESULT 8
US-09-864-761-3705/C
Sequence 3705, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687

```

```

: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Amaxon Sequence Listing Engine vers. 1.1
: SEQ ID NO 3705
: LENGTH: 420
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AP00021.2
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
US-09-864-761-3705

Alignment Scores:
Pred. No.: 2,36e-45 Length: 420
Score: 503.00 Matches: 96
Percent Similarity: 70.29% Conservative: 1
Best Local Similarity: 69.57% Mismatches: 1
Query Match: 19,44% Indels: 40
DB: 10 Gaps: 2

US-09-902-772-4 (1-478) x US-09-864-761-3705 (1-420)

QY 361 TYRAspIysIeuSerATgAlaLeuAArgTYrTYrAspIysAsnIleMetThrIysVal 380
|||
D8 TACGATAAGCTCAGCCGCGCCCTCCGTTACTACTATGACAAGAATCATGACCAAGTCTC 359
|||
QY 381 HisProGluSerSerMetTyrIysTyrProSerAspLeuProTyrMetSerSerTyr 400
|||
D8 359 ----- 359
|||
QY 401 HisGlyIysArgTYrAlaTyrIysPheAspPheHisGlyIleAglAlaLeuGlnPro 420
|||
D8 358 CATGGGAAGCGCTACGCCCTTACAAAGTTCAGACTTCACGGGATCGCCAGGCCCTCCAGCCC 299
|||
QY 420 ----- 420
|||
D8 298 CACCCCGCGGAGTATCTCTTACAAAGTACCCCTCAACACTCCCGGTACATGGGCTCTAT 239
|||
QY 421 HisAlaHisProGlnIysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
|||
D8 238 CACGCCACCCACAGAAGATGAACTTGTGGCGGCCACCCCTCCAGCCCTCCCGGTGACA 179
|||
QY 441 SerSerSerPhePheAlaAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyr 460
|||
D8 178 TCTTCAGATTGTTTGTGTGGCCCCCAACCAATCAGAAATTCACCAACTGGGGGTATATAC 119
|||
QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
|||
D8 118 CCCAACACTAGGCTCCCAACACCAAGCATATGCTTCTATCTGAGCACTTACTAC 65
|||

RESULT 9
US-09-864-761-1653/c
: Sequence 1653, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aeomica-X-1

```

CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 1653
LENGTH: 454
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000163.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
US-09-864-761-1653

Alignment Scores:
Pred. No.: 2,64e-45 Length: 454
Score: 503.00 Matches: 96
Percent Similarity: 70.29% Conservative: 1
Best Local Similarity: 69.57% Mismatches: 1
Query Match: 19.44% Indels: 40
DB: 10 Gaps: 2

US-09-902-772-4 (1-478) x US-09-864-761-1653 (1-454)

QY 361 TyAspLysLeuSerATGAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
|||||
Db 452 TAGATATAGCTCAGCGCGCTCTGCTACTACTATGACAAAGACATCATGACCAAGTC 393
QY 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400

Db 393 ----- 393
QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
|||||
Db 392 CATGGAGAGCGCTACGCGCTACAGTTCGACTTCACAGGAGCGCCAGCGCTCCAGCCC 333
QY 420 ----- 420
Db 332 CACCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGCTCTAT 273
QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProValThr 440
|||||
Db 272 CACGCCACCCACAGAGATGAACTTGTGGCGCCACCTCCAGCCCTCCCGTGACA 213
QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrPheAsnSerProThrGlyLysTyr 460
|||||
Db 212 TCTTCAGTTTTTTTGTGCTGCCCCCAACCCATGAGATTGCACAACTGGGGGTATTATAC 153
QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
|||||
Db 152 CCAAACTAGCTGCTCCACAGCATATGCTTCTATCTGGGCACTTACTAC 99

RESULT 10
US-09-850-799-1
Sequence 1, Application US/09850799
Patent No. US20020090647A1
GENERAL INFORMATION:
APPLICANT: Deneris, Eyan S.
APPLICANT: Eyodoto, Dmitry V.
APPLICANT: Hendricks, Timothy J.
TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
FILE REFERENCE: CASE-03828
CURRENT APPLICATION NUMBER: US/09/850,799
CURRENT FILING DATE: 2001-05-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/360,779
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-26
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1752
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: CDS
LOCATION: (112)..(1131)
US-09-850-799-1

Alignment Scores:
Pred. No.: 4.57e-42 Length: 1752
Score: 481.00 Matches: 122
Percent Similarity: 44.73% Conservative: 18
Best Local Similarity: 38.98% Mismatches: 68
Query Match: 18.59% Indels: 105
DB: 10 Gaps: 10

US-09-902-772-4 (1-478) x US-09-850-799-1 (1-1752)

QY 224 ThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArg 243
|||||
Db 333 ACGGCGGATATGCGCACCCCAAGCGCAAGCTGACCTGCCCTCCGCGCTCCGCC 392
QY 244 IleThrThrArgProAsnProTyrGlnGlnAlaArgSerAlaThrPheHis 263
|||||
Db 393 ATCTCCACCGCCCA ----- GTCGCCGC 416
QY 264 SerHisProThrGlnSerLysAlaThrGlnProSerSerThrValProLysThrGln 283
|||||
Db 417 AGCGATAGACAGAGCGGACACCTCCGCGCTCTGATGACATGACTACAGATTC 476
QY 284 AspGlnArg ----- ProGlnLeuAspProTyrGlnIleLe 295
|||||

```

Db      477 CGTCGAGATGCTCTTTTAAAGNAGGAGAGCCCGAGCTGGGGCCG-----CT 527
QY      295 uGlyProThrSerSerArgLeuAlaAnProGlySerGlyLleGlnLeuTrpGlnPh 315
        |||
Db      528 GAGCCCTGGCGTACAGAAA-----GGCAGCGGGCAGATTCAGTTGTGGCAGTT 575
QY      315 eLeuLeuGluLeuLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSer 335
        |||
Db      576 TCtACTGAGAGCTGCTGGCAGACCGCGGAGACCGCGGTGATCGCGGGAGAGGGGCCCA 635
QY      335 nGlyGluPhLeuMetThrAspProAspGluValAlaArgArgTrpGlyLysLysSe 355
        |||
Db      636 CGCGAGTTCAGGCTCAGCAGCCCGCAGAGGTGGCGCGAGCGCTGGGGGAGGCAAGAG 695
QY      355 rLySProAsnMetAsnTrpAspLysLeuSerArgAlaLeuArgTrpLysLysLysAs 375
        |||
Db      696 CAAGCCCAATGATGAATACAGACAGAGTCAAGTCAAGCTGCGCTGACTACTACGCAAAA 755
QY      375 nileMetThrLysValHisProGlnSerSerMetThrLysTrpProSerAspLeuPr 395
        |||
Db      756 CATCATAGCAAGAGTG----- 771
QY      395 oTyRMetSerSerThrHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyLeuAl 415
        |||
Db      772 -----CACGGCAAGCGCTACGCTACCGCTTGTACTTCCAGGGCCTGGC 815
QY      415 aGlnAlaLeuGlnPro-----HisAlaHisPro----- 424
        |||
Db      816 ACAGGCTTGGCCAGCCACCCCGCGCAGCCGCGCGCTGCCCGCCGCGAGCGGC 875
QY      424 ----- 424
        |||
Db      876 AGCCGCCGCCAGAGATGGCCACTTTACAGCTCCCGCGCTGCTGCTCCACTGCCCTT 935
QY      425 -----GlnLysMetAsnPheValAlaProHisProProAlaLeuProValThrSe 441
        |||
Db      936 CCCGCGCTCTCCAAATCAACCTTATGCGAGCCTGGCGGCGGCGCGCTGCTG 995
QY      441 rSerSerPhePheAlaAlaProAsnProTyTrpTrpAsnSerPro----- 455
        |||
Db      996 CTCT-----TACTGGCTGTGTCACACGCCACGCCGCGCTGC 1031
QY      456 -----ThrGlyLysLeuTrpProAsnThrArgLeuPro----- 466
        |||
Db      1032 CGCCGCCACCGCTGCGCTTACCCACCCCGCGCTTGCAGCCCTCCGCGGCTTTGG 1091
QY      467 -AlaAlaHisMetProSerHisLeuGlyTrpTrpTrpTrpTrpTrpTrpTrpTrp 478
        |||
Db      1092 CGCGGTGGCGCGCTTCGACCTTGGGGGGTCAATTAT 1128

```

```

RESULT 11
US-09-920-300A-1716
; Sequence 1716, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920.300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1716
; LENGTH: 2188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1716
Alignment Scores:

```

```

Pred. No.: 6,48e-38 Length: 2188
Score: 444.00 Matches: 138
Percent Similarity: 38.36% Conservative: 45
Best Local Similarity: 28.93% Mismatches: 134
Query Match: 17.16% Indels: 161
Gaps: 16
US-09-902-772-4 (1-478) x US-09-920-300A-1716 (1-2188)
QY      122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
        |||
Db      491 CACGCGCGCTGGGCAATCCAAAGAACCCCTGCTGGAGTGAGCAACAGATATGCCAG 550
QY      142 TrpLeuGluTrpAlaValLysGlyTrpGlyLeuProAspValAspLysLysLeuPheGln 161
        |||
Db      551 TGGCTTCTTGGGCGCCAAATGAGTTTCACTGTGTGATGATGATGATGATGATGATGATG 607
QY      162 AsnIleAspLysGlyLeuGlyLysMetThrLysAspAspPheGlnArgLeuTrpPro 181
        |||
Db      608 GGCATGATGCGCAGATGCTGTGTAACCTTGGCAGAGACGCTTTCGAGCTGGCACT 667
QY      182 SerTrpAsnAlaAspLysLeuSerHisLeu----- 192
        |||
Db      668 GACTTGTGGGTGACATCTCTGGGAGACATCTGGAGCAAAATGATCAAGAAAACAGAA 727
QY      193 -----HisTrpLeuArg 196
        |||
Db      728 AAGACAGAGATCAATATGAAGAAATTCACACCTCACCTCCGTTCTCATTTGGATTAC 787
QY      197 GluThr----- 198
        |||
Db      788 AGCAATCATTTAGTTTGGCAGAGAGAGCCCTATGGAATGCAGACAGAAATTAC 847
QY      199 -----ProLeu-ProHisLeuThrSerAspAspVa 208
        |||
Db      848 CCCAAAGCGGCTCTCGAGACAGATGTGCCGCTCCACACCCAGCTACTGACCTCT 907
QY      208 LasPlyAlaLeuGlnAsnSerProArgLeu----- 218
        |||
Db      908 GAGCAGGAGTTTCAAGATGTTCCCAAGTCTGGCTCAGCTCCGCTCAGCCTCAGCTACTGC 967
QY      219 -MetHisAlaArgAsnThrGlyLysAlaThrPheIle-----Ph 231
        |||
Db      968 TCtGTCACTGAGAGCTTCCAGGAGGAGCAATTTGCTACCAACAATTCGGGACT 1027
QY      231 eProAsnThrSerValTrpProGluAlaThrGlnArgLysLeuTrpArgProAspLeuPr 251
        |||
Db      1028 CCCAAAGACACGACT---CCCTGAGAGACGGTGGGAGAGCTTCGAGAGCTCAGACTCC 1084
QY      251 oTyRGluln-----AlaArgSerAlaTrp----- 260
        |||
Db      1085 CTCTCCAGTCTGGAGACACCGAGTCTCTCTGTGATGTGACACGGGTTCTCTCTTC 1144
QY      261 -----ThrSerHisSerHisProThrGlnSe 269
        |||
Db      1145 GAGAGCTTCGAAGATGACTGACAGCACTCTCTGCTCAATAAGCAACCACTCTTTC 1204
QY      269 rLySAlaTrpGlnProSerSerSerThrValProLysTrpGluAspGlnArgProGlnLe 289
        |||
Db      1205 AAGGATTACA-----TCCAAAGAGAGAGAGGACCAAGTG-GAGCAAGCAACCAAGTAT 1257
QY      289 uAspProTyRGlulnLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyL 309
        |||
Db      1258 ACCTGACAGCTGTGCTGGCGGCTTCACA-----GGAAGTGAGCC 1296
QY      309 nileGlnLeuTrpGlnPheLeuGlnLeuLeuSerAspSerSerSerSerSerSerSer 329
        |||
Db      1297 TATTACAGCTGGCAGATTTCTCTGAGCTGCTATCAGAAATTCCTGCCAGCTCATTCAT 1356
QY      329 eThrTrpGluGlyThrAsnGlyLysLysMetThrAspProAspGluValAlaArgAr 349
        |||
Db      1357 CAGCTGGAGTGGAGAGAGATGAGATTTAAGTCTGCCGAGCCCGAGTATGAGTGGCGCGCG 1416

```


Db 1595 -GGCGTCAGCCGACGACGAGAGCTGAGTCGCCG-----GACCACCTTGAGCC 1644
 Oy 436 aleuprovalthrserSerSerPhepheaAlaProasnProTyrTrpAsnSerProTh 456
 Db 1645 GGCCCGAGGCTGCTGAGCTAGTGGGAAGCCCATCTGACCAAGCTGCTCCAGAGACCCAG 1704
 Oy 456 rGlyGly-----ThrArgLeuProAlaAlaHisMetProSerHis 473
 Db 1705 GAAGGCGAGATGTGAATATGCCAGAAAGTGGCCAGAAAGACAGTGGCTTATATGCATCC 1764
 Oy 461 oasn-----ThrArgLeuProAlaAlaHisMetProSerHis 473
 Db 1765 CAACACAGGCTCTTGACACAGCTGCTCCTCTGTGGCAGCAGCGCAC 1813
 RESULT 13
 US-09-920-300A-1693
 ; Sequence 1693, Application US/09920300A
 ; Patent No. US20020136728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.547
 ; CURRENT APPLICATION NUMBER: US/09/920,300A
 ; CURRENT FILING DATE: 2001-07-31
 ; NUMBER OF SEQ ID NOS: 1789
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1693
 ; LENGTH: 2268
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-920-300A-1693
 Alignment Scores:
 Pred. No.: 6,83e-38 Length: 2268
 Score: 444.00 Matches: 138
 Percent Similarity: 38.36% Conservative: 45
 Best Local Similarity: 28.93% Mismatches: 134
 Query Match: 17.16% Indels: 161
 Gaps: 16
 US-09-902-772-4 (1-478) x US-09-920-300A-1693 (1-2268)
 Oy 122 GluArgArgValIleValAlaProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
 Db 571 CAGCGGCGGCTGGGCGCATTCGCAAGAACCCCTGCTGTGAGTGAAGCAACAGGTATGCCAG 630
 Oy 142 TripleGluTrpAlaValIleValGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
 Db 631 TGGCTTCTGCGGCGCACCATGAGTCACTGTGTGAGACGTGAATGCGACAGAGGCTC--- 687
 Oy 162 AsnIleAspGlyLysGluLeuGlyMetThrLysAspAspPheGlnArgLeuThrPro 181
 Db 688 GGCATGATGCGCAGATGCTGTGTAACTTGGCAAGAACGCTTTCGAGACTGGACACT 747
 Oy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeu----- 192
 Db 748 GACTTGTGGGTGACATCTCTGTGGAAACATCTGGAGCAAAATGATCAAGAAACCAAGAA 807
 Oy 193 -----HisTyrLeuArg 196
 Db 808 AAGCAGACAGATCAATATGAAGAAATTCACACCTCACCTCGTCTCTCATTTGATTAAC 867
 Oy 197 GluThr----- 198
 Db 868 AGCATATCATATTAGTTTGGCAGAGAGAGCGCCCTATGSAATGACAGACAGAAATTAC 927
 Oy 199 -----ProLeu-ProHisLeuThrSerAspAspVa 208
 Db 928 CCCAAAGCGGCTCTGAGACAGCATGTGTCCGGCTTCACACCACCGTACTCAGCTCT 987

Oy 208 IAspLysAlaLeuGlnAsnSerProArgLeu----- 218
 Db 988 GAGCAGAGGATTTCCAGATTTTCCCAAGCTCTGGCTCCAGCTCCGTCAGGTCACCTACTGC 1047
 Oy 219 -MetHisAlaArgAsnThrGlyAlaThrPheIle-----Ph 231
 Db 1048 TCTGTCACTAGTCAAGCTTCCAGCAGCAACTGTAATTTGTCTACCAACAAATTTGGAGCT 1107
 Oy 231 eProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThrArgProAspLeuPr 251
 Db 1108 CCCAAGACCCAGCACT---CCCTGAGAACGGTGGGCGACCTTCGAGAGCTCAGACTCC 1164
 Oy 251 oTyrGluGln-----AlaArgSerAlaTrp----- 260
 Db 1165 CTCCTCCAGTCTGGAACAGCAGTGTCTTGTCTGTGATGTGCAACAGGCTTCTCTTC 1224
 Oy 261 -----ThrSerHisSerProThrGlnSe 269
 Db 1225 GAGAGCTTCGAAGATGATCGACGACGTCCTCTGCTCATTAAGCCAAACATGCTTTC 1284
 Oy 269 rLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArgProGlnLe 289
 Db 1285 AAGGATTACA-----TCCAAGAGAGAGTACCACAGTG--GAGCAAGGCAACACAGTTAT 1337
 Oy 289 uAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAsnProGlySerGlyG 309
 Db 1338 ACCTGCAGCTGTGCTGGCGGCTTCACA-----GGNAGTGGACC 1376
 Oy 309 nIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuSerAspSerAsnSerAsnGly 329
 Db 1377 TATTCAGCTGTGGAGTTTCTCTGAGACTGCTCATACAGAAATCTCCAGTCACTTCAT 1436
 Oy 329 eThrTrpGluGlyThrAsnGlyLupheLysMetThrAspProAspGluValAlaArgAr 349
 Db 1437 CAGCTGACGTGAACAGCATGAGATTAACTGCGCCAGCCCGATGAGTGGCCCGCG 1496
 Oy 349 gTrrGlyGluArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArgAlaLeuAr 369
 Db 1497 GTGGGGAAGAGGAAATTAAGCCCAAGATTAACCTACAGAGACTGACCGGCGCTTACG 1556
 Oy 369 gTyrTrpTyrAspLysAsnIleMetThrLysValHisProProGluSerSerMetTyrLy 389
 Db 1557 CTACTATTACGACAAAGAACATCATCCACAAAGACG----- 1590
 Oy 389 sTyrProSerAspLeuProTyrMetSerSerTyrHisGlyAspArgTyrAlaTyrLysPh 409
 Db 1591 -----TCGGGGAAGCGCTACGTACCGGCTT 1616
 Oy 409 e-----AspPheHisGlyIleAlaG 416
 Db 1617 CGTGTGAGACTCCAGAACTTCTGGGGTTTACGCCCGGAGAACTGCGACGCCATCTCAG-- 1674
 Oy 416 nAlaLeuGlnProHisAlaHisProGlnLysMetAsnPheValAlaProHisProProAl 436
 Db 1675 -GGCGTCACGCGCAGACGAGGAGCTGAGTGGCGCG-----GACCACCTTGAGCC 1724
 Oy 436 aleuprovalthrserSerSerPhepheaAlaProasnProTyrTrpAsnSerProTh 456
 Db 1725 GGCCCGAGGCTGCTGAGCTAGTGGGAAGCCCATCTGACCAAGCTGCTCCAGAGACCCAG 1784
 Oy 456 rGlyGly-----HisTyrPr 461
 Db 1785 GAAGGCGAGATTTGAATATGCCAGAAAGTGGCCAGAAAGACAGTGGCTTATATGCATCC 1844
 Oy 461 oasn-----ThrArgLeuProAlaAlaHisMetProSerHis 473
 Db 1845 CAACACAGGCTCTTGACACAGGCTGCTCCTCTGTGGCAGACAGCGCAC 1893
 RESULT 14
 US-10-033-528-1693
 ; Sequence 1693, Application US/10033528
 ; Publication No. US20020131971A1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:20:09 ; Search time 3158.2 Seconds

(without alignments)
3805.372 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKKALSVSEDSLE.....IYPNTRLPAMPHSLGTRY 478

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 24791104 segs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09902772/runat_23072003_093658_14926/app_query.fasta.1.1294
-DB=pending.patents.NA_Main -QFMT=fastap -SUFFIX=crpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human40.cdi
-LIST=445 -DOCALLIGN=200 -THR.SCORE=dot -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09902772.ecgn.1.1.4066 -runat_23072003_093658_14926 -NCPD=6 -ICPD=3
-NO_MMAPP -LARGEOUTERY -NEG.SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents_NA_Main:*

1:	/cgn2_6/ptodata/1/pna/US0978.COMB.seq:*
2:	/cgn2_6/ptodata/1/pna/US06.COMB.seq:*
3:	/cgn2_6/ptodata/1/pna/US07.COMB.seq:*
4:	/cgn2_6/ptodata/1/pna/US08.COMB.seq:*
5:	/cgn2_6/ptodata/1/pna/US081.COMB.seq:*
6:	/cgn2_6/ptodata/1/pna/US082.COMB.seq:*
7:	/cgn2_6/ptodata/1/pna/US083.COMB.seq:*
8:	/cgn2_6/ptodata/1/pna/US084.COMB.seq:*
9:	/cgn2_6/ptodata/1/pna/US085.COMB.seq:*
10:	/cgn2_6/ptodata/1/pna/US086.COMB.seq:*
11:	/cgn2_6/ptodata/1/pna/US087.COMB.seq:*
12:	/cgn2_6/ptodata/1/pna/US088.COMB.seq:*
13:	/cgn2_6/ptodata/1/pna/US089.COMB.seq:*
14:	/cgn2_6/ptodata/1/pna/US090.COMB.seq:*
15:	/cgn2_6/ptodata/1/pna/US091.COMB.seq:*
16:	/cgn2_6/ptodata/1/pna/US092.COMB.seq:*
17:	/cgn2_6/ptodata/1/pna/US093.COMB.seq:*
18:	/cgn2_6/ptodata/1/pna/US094.COMB.seq:*
19:	/cgn2_6/ptodata/1/pna/US095.COMB.seq:*
20:	/cgn2_6/ptodata/1/pna/US095B.COMB.seq:*
21:	/cgn2_6/ptodata/1/pna/US095C.COMB.seq:*
22:	/cgn2_6/ptodata/1/pna/US095D.COMB.seq:*
23:	/cgn2_6/ptodata/1/pna/US096.COMB.seq:*
24:	/cgn2_6/ptodata/1/pna/US096B.COMB.seq:*
25:	/cgn2_6/ptodata/1/pna/US096C.COMB.seq:*
26:	/cgn2_6/ptodata/1/pna/US096D.COMB.seq:*
27:	/cgn2_6/ptodata/1/pna/US096E.COMB.seq:*
28:	/cgn2_6/ptodata/1/pna/US097A.COMB.seq:*

29:	/cgn2_6/ptodata/1/pna/US097B.COMB.seq:*
30:	/cgn2_6/ptodata/1/pna/US097C.COMB.seq:*
31:	/cgn2_6/ptodata/1/pna/US098A.COMB.seq:*
32:	/cgn2_6/ptodata/1/pna/US098B.COMB.seq:*
33:	/cgn2_6/ptodata/1/pna/US098C.COMB.seq:*
34:	/cgn2_6/ptodata/1/pna/US099A.COMB.seq:*
35:	/cgn2_6/ptodata/1/pna/US099B.COMB.seq:*
36:	/cgn2_6/ptodata/1/pna/US099C.COMB.seq:*
37:	/cgn2_6/ptodata/1/pna/US099D.COMB.seq:*
38:	/cgn2_6/ptodata/1/pna/US100A.COMB.seq:*
39:	/cgn2_6/ptodata/1/pna/US100B.COMB.seq:*
40:	/cgn2_6/ptodata/1/pna/US101A.COMB.seq:*
41:	/cgn2_6/ptodata/1/pna/US101B.COMB.seq:*
42:	/cgn2_6/ptodata/1/pna/US102A.COMB.seq:*
43:	/cgn2_6/ptodata/1/pna/US102B.COMB.seq:*
44:	/cgn2_6/ptodata/1/pna/US6000.COMB.seq:*
45:	/cgn2_6/ptodata/1/pna/US6001.COMB.seq:*
46:	/cgn2_6/ptodata/1/pna/US6002.COMB.seq:*
47:	/cgn2_6/ptodata/1/pna/US6003.COMB.seq:*
48:	/cgn2_6/ptodata/1/pna/US6004.COMB.seq:*
49:	/cgn2_6/ptodata/1/pna/US6005.COMB.seq:*
50:	/cgn2_6/ptodata/1/pna/US6006.COMB.seq:*
51:	/cgn2_6/ptodata/1/pna/US6007.COMB.seq:*
52:	/cgn2_6/ptodata/1/pna/US6008.COMB.seq:*
53:	/cgn2_6/ptodata/1/pna/US6009.COMB.seq:*
54:	/cgn2_6/ptodata/1/pna/US6010.COMB.seq:*
55:	/cgn2_6/ptodata/1/pna/US6011.COMB.seq:*
56:	/cgn2_6/ptodata/1/pna/US6012.COMB.seq:*
57:	/cgn2_6/ptodata/1/pna/US6013.COMB.seq:*
58:	/cgn2_6/ptodata/1/pna/US6014.COMB.seq:*
59:	/cgn2_6/ptodata/1/pna/US6015.COMB.seq:*
60:	/cgn2_6/ptodata/1/pna/US6016.COMB.seq:*
61:	/cgn2_6/ptodata/1/pna/US6017.COMB.seq:*
62:	/cgn2_6/ptodata/1/pna/US6018.COMB.seq:*
63:	/cgn2_6/ptodata/1/pna/US6019.COMB.seq:*
64:	/cgn2_6/ptodata/1/pna/US6020.COMB.seq:*
65:	/cgn2_6/ptodata/1/pna/US6021.COMB.seq:*
66:	/cgn2_6/ptodata/1/pna/US6022.COMB.seq:*
67:	/cgn2_6/ptodata/1/pna/US6023.COMB.seq:*
68:	/cgn2_6/ptodata/1/pna/US6024.COMB.seq:*
69:	/cgn2_6/ptodata/1/pna/US6025.COMB.seq:*
70:	/cgn2_6/ptodata/1/pna/US6026.COMB.seq:*
71:	/cgn2_6/ptodata/1/pna/US6027.COMB.seq:*
72:	/cgn2_6/ptodata/1/pna/US6028.COMB.seq:*
73:	/cgn2_6/ptodata/1/pna/US6029.COMB.seq:*
74:	/cgn2_6/ptodata/1/pna/US6030.COMB.seq:*
75:	/cgn2_6/ptodata/1/pna/US6031.COMB.seq:*
76:	/cgn2_6/ptodata/1/pna/US6032.COMB.seq:*
77:	/cgn2_6/ptodata/1/pna/US6033.COMB.seq:*
78:	/cgn2_6/ptodata/1/pna/US6034.COMB.seq:*
79:	/cgn2_6/ptodata/1/pna/US6035.COMB.seq:*
80:	/cgn2_6/ptodata/1/pna/US6036.COMB.seq:*
81:	/cgn2_6/ptodata/1/pna/US6037.COMB.seq:*
82:	/cgn2_6/ptodata/1/pna/US6038.COMB.seq:*
83:	/cgn2_6/ptodata/1/pna/US6039.COMB.seq:*
84:	/cgn2_6/ptodata/1/pna/US6040.COMB.seq:*
85:	/cgn2_6/ptodata/1/pna/US6041.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2436	94.1	1528	34	US-09-902-772-3
2	2262.5	87.4	1447	34	US-09-902-772-1
3	2195.5	84.8	3309	25	US-09-652-109-9617
4	2195.5	84.8	3309	28	US-09-716-953-1936
5	2195.5	84.8	3309	29	US-09-721-589-5498
6	2195.5	84.8	3309	29	US-09-726-806-4625

```

7 2195.5 84.8 3309 29 US-09-726-811-4920 Sequence 4920, Ap
8 2180.5 84.3 1389 35 US-09-948-941-233 Sequence 233, App
9 2180.5 84.3 3166 1 PCT-US02-23913-98 Sequence 98, App1
10 2180.5 84.3 3166 30 US-09-784-356-26 Sequence 26, App1
11 2180.5 84.3 3166 38 US-10-021-660-26 Sequence 26, App1
12 2180.5 84.3 3166 42 US-10-205-823-98 Sequence 26, App1
13 2017.5 78.0 2172 80 US-60-360-207-355 Sequence 355, App
14 1808.5 69.9 3126 18 US-09-442-384A-475 Sequence 475, App
15 1808.5 69.9 3126 18 US-09-442-384B-475 Sequence 475, App
16 1482.5 57.3 1359 38 US-10-035-832-1210 Sequence 1210, App
17 1482.5 57.3 1359 38 US-10-052-482-198 Sequence 198, App
18 1482.5 57.3 2938 7 US-08-343-443-3 Sequence 3, App11
19 1482.5 57.3 2957 1 PCT-US02-18947-710 Sequence 710, App
20 1482.5 57.3 2957 38 US-10-007-926A-295 Sequence 295, App
21 1482.5 57.3 2957 38 US-10-035-832-1209 Sequence 1209, App
22 1482.5 57.3 2957 39 US-10-052-482-197 Sequence 197, App
23 1482.5 57.3 2957 41 US-10-172-118-710 Sequence 710, App
24 1482.5 57.3 3298 25 US-09-652-814-9467 Sequence 9467, App
25 1482.5 57.3 3298 27 US-09-699-899-6013 Sequence 6013, App
26 1482.5 57.3 3298 28 US-09-716-953-1954 Sequence 1954, App
27 1482.5 57.3 3298 28 US-09-721-589-5554 Sequence 1139, App
28 1482.5 57.3 3298 29 US-09-726-171-2148 Sequence 2148, App
29 1482.5 57.3 3298 29 US-09-726-806-4681 Sequence 4681, App
30 1482.5 57.3 3298 17 US-09-396-870-7615 Sequence 7615, App
31 1482.5 57.3 3372 17 US-10-035-832-1207 Sequence 1207, App
32 1474.5 57.0 1359 38 US-10-052-482-195 Sequence 195, App
33 1474.5 57.0 1359 38 US-09-053-375B-627 Sequence 627, App
34 1474.5 57.0 1729 14 US-10-035-832-1206 Sequence 1206, App
35 1474.5 57.0 1729 39 US-10-052-482-194 Sequence 194, App
36 1474.5 57.0 1729 80 US-60-360-207-10845 Sequence 10845, App
37 1474.5 57.0 1729 76 US-60-324-185-33223 Sequence 33223, App
38 1469.5 56.8 3110 76 US-60-172-373-5267 Sequence 2567, App
39 1468.5 56.7 1307 61 PCT-US02-25766-2596 Sequence 2596, App
40 1394.5 53.9 1932 14 US-09-053-375B-327 Sequence 327, App
41 1394.5 53.9 1932 18 US-09-442-384A-530 Sequence 530, App
42 1394.5 53.9 1932 18 US-09-442-384B-530 Sequence 530, App
43 1394.5 53.9 1932 18 US-09-726-806-3434 Sequence 3434, App
44 1061.0 41.0 661 29 US-09-726-806-3434 Sequence 7574, App
45 1019.0 39.4 840 25 US-09-652-109-7574 Sequence 7574, App

```

ALIGNMENTS

```

RESULT 1
US-09-902-772-3
: Sequence 3, Application US/099022772
: GENERAL INFORMATION:
: APPLICANT: Chugai Selyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: FILE REFERENCE: chugai selyaku kabushiki kaisha 5001
: CURRENT APPLICATION NUMBER: US/09/902,772
: PRIOR FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/08/878,177
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1528
: TYPE: DNA
: ORGANISM: C-ery gene, chicken DNA
US-09-902-772-3

Alignment Scores:
Pred. No.: 4,41e-178 Length: 1528
Score: 2436.00 Matches: 458
Percent Similarity: 91.97% Conserves: 0
Best Local Similarity: 91.97% Mismatches: 0
Query Match: 94.13% Indels: 40
DB: 34 Gaps: 2

US-09-902-772-4 (1-478) x US-09-902-772-3 (1-1528)

```

```

Qy 1 MetalaserthrilelysgluAlaleuSerValValSerluaspglInSerleupheglu 20
Db 63 ATGCAGACACTATTAAAGAGCATTTATCACTGTGATGTAAGACCAAGCTCTTTGAG 122
Qy 21 CysAlatylrglySerProHisleuAlalysThrclunetHralasSerSerSerGlu 40
Db 123 TGTCTCCACGAGATGCCGCCACCTTGCACAAAGCAAAATGACAGGCTCTCTTCCAGTAA 182
Qy 41 TyrlcglInThrSerlysmetSerProArvalProglInlnaSPTrpleuSeglPro 60
Db 183 TATGGCAACATCAAAAGATGAGCCGCCGCTTCCACGACGAGCTGTATCAACGCC 242
Qy 61 ProAlaargValThrilelysmetClyCysAnProAsnInValaAnglySerArasn 80
Db 243 CCGGCCAGAGTTACCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
Qy 81 SerProaspaspCysSerSerValAlalysGlyClylysmetValSerSerSerAspVal 100
Db 303 TCACCTGATGCTCAGCTGAGGCAAAAGAGGAAATGTTAGCACTTACAGACAAATGTT 362
Qy 101 GlyMetasnyrglySerTyrmecglulysHisIlepropProAsnMetThrThr 120
Db 363 GGGATGACACTATGAGACTACATGAGAGAGAGATATTCGCCCTCCAAATATGACAAAC 422
Qy 121 AsncluarArgValIleValProAlasProThrleutPserThrAspHisValArg 140
Db 423 AATCAAGCAAGATTAATGTCGCCAGCATCTCTGATGAGACAGACAGCAATGTCGG 482
Qy 141 GlntPleuglntPralaVallysgluTyltyglleuProaspValaspIleleuPhe 160
Db 483 CAGTGCTGAGTGGAGGAGTGAAGAGATATGCTTTCAGACGTGACATCTTGTTC 542
Qy 161 GlnaSnIleaspgllysgluLeuCyslysmetThrlyspaspPheglInArgLeuThr 180
Db 543 CAGAACATTTGATGGAAAGATTTGTATAAATGACCAAAATGACTTCCAGAGCTCAGC 602
Qy 181 ProSerlyrAsnAlaaspIleleuSerHisleuHstlyrleuArggluThrProleu 200
Db 603 CCGAGCTATACGAGATATCTCTCTGACACCTTACATCAACCAAGAGACATCTCT 662
Qy 201 ProHisleuThrSerAspAspValaspIlyslaleuGlInasSerProArgleuMetHis 220
Db 663 CCACATTTGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
Qy 221 AlaArasnThrlysglyAlalathrPheIlePheProAsnThrSerValTyrlProglua 240
Db 723 GCTAGAAACACAGAGAGAGGACACTTTATTTTCCAAATACATCAGATTCCACAGAGCA 782
Qy 241 ThrGlnArgIleThrThrArgProaspLeuProTyrlglInAlaArgSerAlatrp 260
Db 783 ACGCAAGAATTAACACAGAGCCAGATTACCTTATGACGAGAGAGAGATCAGCTGG 842
Qy 261 ThrSerHisSerHisProThrInSerlyserValAlathrGlnProSerSerThrValPro 280
Db 843 ACGAGTACAGCCATCCACCTCACTCAAAAGCTTACCAACCATCATCTTCAACAGTCCC 902
Qy 281 LysThrGlnuspglInatProglInleuaspProTyrlglInleuenglyProThrSerSer 300
Db 903 AAAACAGACAGACGAGCTCTCACTGATGATGATGATGATGATGATGATGATGATGATG 362
Qy 301 ArgLeuAlaasnProgllyserGlyglInleuThrPclInPheleuLeuGluLeu 320
Db 963 CGTCTTCCAAATCCAGGAGTGGGACATACAGTATGAGGACAGTCCACTACAGGCTCTG 1022
Qy 321 SerAspSerSerAsnSerArasnCysIleThrTrpleuInlyThrAsnlyglupheLyMet 340
Db 1023 TCGGACAGCTCCAACTCAATGATCATCTGGAGGAGGCAAAATGGGAGTTCCAAATG 1082
Qy 341 ThrAspProaspglValAlaArgArgTrpGlygluArglySerSerlyspAsnMetAsn 360
Db 1083 ACAGACCCCTATGAAGTGGCTCGGCGTTGGGAGAGAGAGAAACCAACCTACATGAC 1142
Qy 361 TyrAspIlyleuSerArGalaleuArgTyrlTyrlAspIlyAsnIleMetThrlyVal 380

```

```

|||||
Db 1143 TATGCAAACTGAGCGGCGACTGCTACTACTATGACAAAATATATGACTTAAGTT 1202
QY 380 -----
Db 1203 CATGGTAAACGCTATGCTTACAAATTTGATTTCCACGGAATCGCTCAGCCCTCCAGCCT 380
QY HisProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 381 HisProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1263 CACCTCCAGATCATTCATGACAAATACCATGACGCTCCCTACATAGATGCTTAC 1322
QY HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1322 -----
QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1323 CATGCACACCCCAAGAAATGAACTTTTGTAGCTCCCATCCCTGCTTGGCCGTAC 1382
QY SerSerSerPhePheAlaAlaProAsnProTyrTyrPaspSerProThrGlyGlyTyr 460
Db 1383 TCATCCAGCTTTTGGCTGCCCCCTAATCCATGAAATGCACCACTGGAGGCAATTCAC 1442
QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1443 CCCAATACGAGGCTGCCAGCTGCTCATATGCTTCCCTTGGCAGCTACTAC 1496

```

RESULT 2

```

US-09-902-772-1
: Sequence 1, Application US/09902772
: GENERAL INFORMATION:
: APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
: TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
: FILE REFERENCE: Chugai seiyaku kabushiki kaisha 5001
: CURRENT APPLICATION NUMBER: US/09/902,772
: PRIOR FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US/08/878,177
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1447
: TYPE: DNA
: ORGANISM: C-11 gene, c-ery gene w/ deletion, chicken DNA
US-09-902-772-1

```

Alignment Scores:

```

Pred. No.: 1,07e-164 Length: 1447
Score: 2262.50 Matches: 430
Percent Similarity: 86.35% Conservative: 0
Best Local Similarity: 86.35% Mismatches: 1
Query Match: 87.42% Indels: 67
DB: 34 Gaps: 3

```

US-09-902-772-4 (1-478) x US-09-902-772-1 (1-1447)

```

QY 1 MetAlaSerThrIleLysGlnAlaLeuSerValValSerGluAspGlnSerLeuPheGln 20
Db 63 ATGGCAACGACTAATTAAAGAGACATTTACAGTGGTGAAGAACACAGCTCTGTTTGAG 122
QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGlnMetThrAlaSerSerSerGln 40
Db 123 TGTGCTTAGCGATGCGCCACCTTGCAAGACAGAAATGACAGCCCTCTTCCAGTAA 182
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
Db 183 TATGGGCAACATCAAAAGATGAGCGCGCTTCCCGCAGGAGCTGTTATCAGAGCC 242
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db 243 CCGGCCAGAGTTACATTAAAGATGAGTGTAAACCAACAGGTAAATGAGGTCAAGGAAAT 302

```

```

QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db 303 TCACCTGATGACTGACGCGGTGGCAAGAGAGGAAATGGTTAGCAGTTTCAGACAAATGTT 362
QY 101 GlyMetAsnTyrGlySerTyrMetGlnGlnLysHisIleProProProAsnMetThrThr 120
Db 363 GGATGAACTATGGAAGCTACATGGAAGAAGAGCATATCCGCCCTCCAAATATGACAAAC 422
QY 121 AsnGluArgValIleValAlaProAlaAspProThrLeuTyrPserThrAspHisValArg 140
Db 423 AATGACGAAAGATTAATGTGCTCCAGCAATCTTACGTTATGAGAGCAGACCATGTACGG 482
QY 141 GlnThrLeuGlnUTRPalValLysGlnLysTyrLeuProAspValAspIleLeuLeuPhe 160
Db 483 CAGTGGCTGGAGTGGGCGAGTGAAGATGATGCTCTCCAGACGTGACATCTTGTGTTTC 542
QY 161 GlnAsnIleAspGlyLysGlnLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db 543 CAGAACTATGATGGAAAGAGTTGTGTAATGACCAAAAGATGACTTCCAGAGCTCAG 602
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGlnThrProLeu 200
Db 603 CCGAGCTTAACGAGATATCTCTCTGACACCTACCTACCTCAGAGAGA----- 656
QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
Db 656 -----
QY 221 AlaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGlnAla 240
Db 657 -----GGAGCCACTTTTATTTTCCAAATACATACAGTTTACCAGAAACA 701
QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgSerAlaTyr 260
Db 702 ACGCAAAAGATTAACAAACAGGAGATTTACCTTATGAGCAAGCAGAGAGATTCAGCCTGG 761
QY 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 762 ACGAGTCAACGCCATCCAGCTCAGTCAAAACCTCCCAACATCATCTTCAACAGTGGCC 821
QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db 822 AAAACAGAAAGACAGCGCTCCAGTATGATCTTATGATCTTGGACGACGACGACG 881
QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeuLeu 320
Db 882 CGTCTTGCAAAATCCAGGAGTGGGAGATACAGCTATGAGCAGTTCCTACTGAGCTCTG 941
QY 321 SerAspSerSerAsnSerAsnCysIleThrThrProGlnGlyThrAsnGlyGlnPheLysMet 340
Db 942 TCGGACAGCTCACTCCACCTGCAATCAGTCACTGGGAGGAGCAAAATGGGAGTTCAAGATG 1001
QY 341 ThrAspProAspGlnValAlaArgArgTyrProGlnArgLysSerLysProAsnMetAsn 360
Db 1002 ACAGACCTGATGAAGTGGCTCGGCGTTGGGAGAGAGAGAAACCAACCTAACATGAGAC 1061
QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db 1062 TATGACAAACTCAGCCGCTGCACTTCGCTACTACTATGACAAAAATATATGACTTAAAGTT 1121
QY 380 -----
Db 1122 CATGGTAAACGCTATGCTTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAGCCT 1181
QY 381 HisProProgluserSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1182 CACCTCCAGATCATCATCATATACAAATACCAATACCAAGCTCCGCTCATATGATGCTTAC 1241
QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1241 -----
QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440

```

Db 1242 CATCACACCCCGAGAGTAACTTTGACTCCCTCCCTGCTTTCCCGTAACC 1301
 Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyLeuTyr 460
 Db 1302 TCATCCAGCTTTTGGTGGCCCTTAATCCACTGGAATTCACCAACTGGAGGACTTAC 1361
 Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
 Db 1362 CCCAATACAGGCTGCCAGCTGCTCATATGCTTCCTCCACTTGGACCTACTAC 1415

RESULT 3

US-09-652-109-9617
 ; Sequence 9617, Application US/09652109
 ; GENERAL INFORMATION:
 ; APPLICANT: McCarthy, Sean A.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 1600,1180-001
 ; CURRENT APPLICATION NUMBER: US/09/652,109
 ; PRIOR FILING DATE: 2000-08-30
 ; PRIOR APPLICATION NUMBER: 60/151,128
 ; PRIOR FILING DATE: 1999-08-30
 ; NUMBER OF SEQ ID NOS: 10105
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9617
 ; LENGTH: 3309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-652-109-9617

Alignment Scores:

Pred. No.: 4,4e-159 Length: 3309
 Score: 2195.50 Matches: 416
 Percent Similarity: 84.97% Conservative: 8
 Best Local Similarity: 83.37% Mismatches: 10
 Query Match: 84.83% Indels: 65
 Db: 25 Gaps: 4

US-09-902-772-4 (1-478) x US-09-652-109-9617 (1-3309)

Oy 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 128 ATGGCCAGCACTATTAAAGGAGCCTTATCACTTGTGATGAGGACCACTCTTTTGTGAG 187
 Oy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40
 Db 188 TGTGCTACGAGCAAGCCACACCTGGCTAAGACAGATGACCGGCTCTCTCCAGCCAC 247
 Oy 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 Db 248 TATGACAGACTTCCAGATGAGCCACGCTCCCTCAGCAGATGGCTGTCTCAACCC 307
 Oy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 308 CCACGCCAGGCTCACATCAAAATGTAACCTTCCCGCAGGTGAATGGCTCAAGAAC 367
 Oy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 Db 368 TCTCTGATGATGACTGTGTGGCCAAAGCGGGAAGATGTGGGACACCCAGACACCTT 427
 Oy 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
 Db 428 GGGATGAACACGCGACACTACATGAGAGAACACATGACACCCCAACATGACACAG 487
 Oy 121 AsnGlnArgArgValIleValProAlaAspProThrIleuTrpSerThrAspHisValArg 140
 Db 488 AACAGGCGCAGAGTATTCGCGCAGAGATCTTACGCTATGAGATAGACACATGCTCGG 547
 Oy 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
 Db 548 CAGTGGCTGAGTGGGCGGTGAAGAATATGGCTTCCACAGCTCAACATCTTTTATTC 607

Oy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 608 CAGAACATCGATGGAGGAACTGTGCAAGATGACCAAGAGCAGCTTCCAGAGGCTCAC 667
 Oy 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluThrProLeu 200
 Db 668 CCCAGCTAACAGCGCGATCTCTCTCATATCTCCACATCCACAGACACACCTCTT 727
 Oy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 Db 728 CCACATTTGACTTGAAGTATGATGTGATTAACCTTACAAACTCTCCACGGTTATATCAT 787
 Oy 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 Db 788 GCTAGAAAACACA----- 799
 Oy 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrp 260
 Db 800 -----GATTTACCATATGAGCCCCCAGAGAGATCAGCTGG 835
 Oy 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
 Db 836 ACCGGTCACGGGCCACCCACAGCCCGCAGTCCGAAGCTGOTCAACCATCTCTCCACAGTG 895
 Oy 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
 Db 896 CCCAAACTGAAAGCACAGCTCTCTCATATGATTCATGATTTCTGGACCAACAGT 955
 Oy 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGlnLeu 319
 Db 956 AGCCGCTTGCAAAATCCAGGAGTGCGCAGATCCAGCTTGGAGATTCTCTCGAGCTC 1015
 Oy 320 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLysPheLys 339
 Db 1016 CTGTGGAGACAGCTCAACTCAGCTGATCACTCGGGAAGGACCAACAGGGAGTTCAAG 1075
 Oy 340 MetThrAspProAspGluValAlaArgArgTrpGlyLysLysSerLysProAsnMet 359
 Db 1076 ATGACGATCCCGACGAGGTGGCGCGCGCTGGGAGAGCGGAAGACCAACCCACATG 1135
 Oy 360 AspTyrAspLysLeuSerArgAlaLeuArgTyrTrpTyrAspLysAsnIleMetThrLys 379
 Db 1136 AACTACGATTAAGCTCAGCGCGGCTCGTACTACATGATGACAAAGATCATGACCAAG 1195
 Oy 380 ValHisProProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399
 Db 1196 GTC----- 1198
 Oy 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
 Db 1199 ---CATGGGAAGCGCTACGGCTTCAAGTTGCACTTCCACGGGATGGCCAGGCCCTCAG 1255
 Oy 420 Pro----- 420
 Db 1256 CCCACCCCGGAGTATCTGTACAGATACCCCTCAGACCTCCCGTACATGGGCTCC 1315
 Oy 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
 Db 1316 TATACGCCCAACAGAGAAAGATGAATTTGTGGCGCCACCCCTCAGGCCCTCCCGGTG 1375
 Oy 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIle 459
 Db 1376 ACATCTTCCAGTTTGTGTCGCCCAACCCCAATCTGGAATTCACCAACTGGGGGTATA 1435
 Oy 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 Db 1436 TACCCCAACACTAGGCTCCACACGCAATATGCTTCTCATCTGCGGACTTACTAC 1492

RESULT 4

US-09-716-953-1936
 ; Sequence 1936, Application US/09716953
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.

```

; APPLICANT: Holtzman, Douglas A.
; APPLICANT: McCarthy, Sean
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2044-001
; CURRENT APPLICATION NUMBER: US/09/716,953
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/167,413
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 2620
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1936
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-716-953-1936

Alignment Scores:
Pred. No.: 4,4e-159      Length: 3309
Score: 2195.50          Matches: 416
Percent Similarity: 84.978      Conservative: 8
Best Local Similarity: 83.37%      Mismatches: 10
Query Match: 84.83%      Indels: 65
DB: 28      Gaps: 4

US-09-902-772-4 (1-478) x US-09-716-953-1936 (1-3309)

OY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 128 ATGGACACACTATATTAAGAGAGCCCTTATCAGTTGGATGAGAGACCACTGCTTTGAG 187
OY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerL 40
DB 188 TGTGCTTGGAGACCCACCTGGCTGAAGACAGATGACCGCTCTCTCCAGCGAC 247
OY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnPro 60
DB 248 TATGACAGACTCTCAATGATGAGCCAGCGCTCCCTGACGAGATGGCTGTCTCAACCC 307
OY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAGCCAGGCTGACCATCAATGGAATGTAACCTTACCCAGTGAATGGCTCAAGAAC 367
OY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerSerAspAsnVal 100
DB 368 TCTCTGTGATGATGAGTGTGGCCAAAGCGGAGATGTGGCGACGACACCGCTT 427
OY 101 GlyMetAsnTyrGlySerTyrMetGlnGluLysHisIleProProProAsnMetThrThr 120
DB 428 GGGATGACACTACGAGCTACATGAGAGAGACATGCCACCCCAACATGACACAG 487
OY 121 AsnGluArgValIleValProAlaAspProThrLeuTyrPsrThrAspHisValArg 140
DB 488 AACGAGCCAGAGTATCTGCCAGACATCTTACGCTATGAGATACAGACATGTGGCG 547
OY 141 GlnTrpLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
DB 548 CAGTGGCTGGAGTGGCGGTGAAGAATATGCGCTCCAGACGCTAACATCTTGTATTTC 607
OY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB 608 CAGAAACATCGATGGAGAGAACTGTGCAAGATGACCAAGACGACTTCCAGAGCTCAC 667
OY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB 668 CCCAGCTTCAACAGCGCGACATCTCTCTCACATCTCCACTCCAGAGAGACATCTCTT 727
OY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB 728 CCACATTTGACTTTCGATGATGTGATTAAGCCTTACAAACTCTCCACGGGTTAATGAT 787
OY 221 AlaArgAsnThrGlyLysValaThrPheIlePheProAsnThrSerValTyrProGluAla 240

```

```

DB 788 GCTAGAACACA----- 799
OY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGlnAlaArgArgSerAlaTrp 260
DB 800 -----GATTTACATATGAGACCCCGACAGAGATACAGCTTG 835
OY 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
DB 836 ACCGGTCAAGGCGACCCAGCCAGCGAGTGAAGCTGCTCAACCTCTTCCACAGATG 895
OY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB 896 CCCAAATCTGAAGACACCGCTCTCAGTATGATCTTATCAGATTTCTTGACCAACAGT 955
OY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPheLeuLeuGluLeu 319
DB 956 AGCGGCTTGCATAATCCAGAGCAGTGGCAGATCCAGCTTGGCAGATCTCTCTGAGCTC 1015
OY 320 LeuSerAspSerSerAsnSerAsnGlySerThrTrpGluGlyThrAsnGlyGluPheLys 339
DB 1016 CTGTGGACACTCTCACTCACTGATCACTGGAAGGACCAACGGGAGATTCAAG 1075
OY 340 MetThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMet 359
DB 1076 ATGACAGATCCCGACAGAGAGTGGCGCGCTGGGAGAGCGGAAGACCAACCAACATG 1135
OY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB 1136 AACTACGATTAAGCTACCGCGCGCTCCCTTACTACTATGACAAACATCATGACCAAG 1195
OY 380 ValHisProProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399
DB 1196 GTC----- 1198
OY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB 1199 ---CATGGAGCGGTAGCGCTACAGATGTGCACTTCCACGGGATGCGCCCTCCAG 1255
OY 420 Pro----- 420
DB 1256 CCCCACCCCGGAGTCACTCTGTACAAAGTACCCCTCAGACCTCCCTACATGGGCTCC 1315
OY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProAl 439
DB 1316 TATCAGCGCCACCCACAGAAAGTGAATTTGTGGGCGCCACCTCCAGCCCTCCCTG 1375
OY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPAsnSerProThrGlyLys 459
DB 1376 ACATCTTCCAGTTTTTTGGTGGCCCAAAACCATATCTGGAATTCACCAATGGGGGTATA 1435
OY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
DB 1436 TACCCCAACACTAGCTCCACAGCATATGCTTCTCACTGTGGGCACTTACTAC 1492

RESULT 5
US-09-721-589-5498
; Sequence 5498, Application US/09721589
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Villereal, Jean-Luc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2045-001
; CURRENT APPLICATION NUMBER: US/09/721,589
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/167,380
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 7017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5498
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-721-589-5498

Alignment Scores:

Pred. No.: 4.4e-159 Length: 3309
 Score: 2195.50 Matches: 416
 Percent Similarity: 84.97% Conservative: 8
 Best Local Similarity: 83.37% Mismatches: 10
 Query Match: 84.83% Indels: 65
 DB: 29 Gaps: 4

US-09-902-772-4 (1-478) x US-09-721-589-5498 (1-3309)

```

OY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 128 ATGGCCACACTATTAAAGGAAGCCCTTATCAGTTGTGATGAGGACCACTCGCTTTGAG 187
OY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB 188 TGTGCTTACGAGACCCACACTGCTTACAGACAGATGACCGCTCTCTCCAGCCAG 247
OY 41 TyrGlnGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
DB 248 TATGACAGACTTCCAAAGATGAGCCACGCGCTCCAGACAGATTGGCTGTCAACCC 307
OY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAGCCAGGCTCACCATATAATGGAATGTAACCTACCCAGGTGAATGGCTCAAGAAC 367
OY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB 368 TCTCTGTATGATGACAGTGTGCGCCAAAGCGGAGAGATGGTGGCCAGCCAGACCGCT 427
OY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB 428 GGGATGACATCAGCGAGCTACATGAGAGAGAACATGCCACCCCAACATGACACAG 487
OY 121 AsnGluArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
DB 488 AACGAGCCCAAGTATGTCGCCACAGATCCTACGCTATGAGACAGACCATGTGGCG 547
OY 141 GlnTrpLeuGluTrpAlaValLysGluTrpGlyLeuProAspValAspIleLeuLeuPhe 160
DB 548 CAGTGGCTGGAGTGGCGGCTCAAAAGATATGGCTTCCAGACGTCACATCTTGTATTC 607
OY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB 608 CAGAACATCGATGGAGAGAACTGTCAAAGATGACCAAGACGACTTCCAGAGGCTCAC 667
OY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB 668 CCCAGCTACAAAGCCGACATCTCTCTCCACATCTCCACTCCCTCAGAGAGACTCTCTT 727
OY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB 728 CCACATTTGACTTCCAGATGATGTGATTAAGCTTACAAACTCTCCAGGTTAATGCAT 787
OY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
DB 788 GCTAGAAACACA----- 799
OY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260
DB 800 -----GATTACCATATGAGCCGCCAGAGATCAACCCGCG 835
OY 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
DB 836 ACCGCTCAGCGCCACCCACGCCCAAGTCTCAACCATCTCTCTCCACACATG 895
OY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB 896 CCCAAACTGAAGACACAGCTCTCAGTTAGATCTTATCAGATTCCTTGAGCAACAGAGT 955
OY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 319
  
```

```

DB 956 AGCCGCTTGCAGAAATCCAGGCAAGTGGCCAGATCCAGCTTTGGCACTTCTCTGGAGCTC 1015
OY 320 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLupheLys 339
DB 1016 CTGTGGACAGCTCCAACTCCAGCTGCATTCACCTTGGGAAGCACCACAGGGAGTTCAAG 1075
OY 340 MetThrAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 359
DB 1076 ATGACGGATCCCGACAGAGAGTGGCCGCGCTGGGAGAGAGCGGAAGAACCCAAACATG 1135
OY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB 1136 AACTAGATTAAGCTCAGCGCGCCCTCCGTTACTATGACAAAGAACATCATGACCAAG 1195
OY 380 ValHisProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSer 399
DB 1196 GTC----- 1198
OY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB 1199 ---CATGGGAAGCGCTACGCTACAAAGTTGCACCTTCCACGGGATGCCAGGCCCTCAG 1255
OY 420 Pro----- 420
DB 1256 CCCACCCCGGAGTCACTCTGTACAAGTACCCCTAGACCTTCCGTACATGGGCTCC 1315
OY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProAlaLeuProVal 439
DB 1316 TATCAGCGCCACCCACAGATGACTTTGTGGGCGCCACCCACCTCCAGCCCTCCG 1375
OY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysIle 459
DB 1376 ACATCTTCAGATTTTGTGTGTCGCCCAAAACCATCTGGAATTCACAACTGGGGGTATA 1435
OY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTrp 478
DB 1436 TACCCCAACTAGAGCTCCACCAAGCATATGCTTCATCTGGGACCTTACTAC 1492
  
```

RESULT 6
 US-09-726-806-4625
 ; Sequence 4625, Application US/09726806
 ; GENERAL INFORMATION:
 ; APPLICANT: Galvin, Katherine
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 1600.2028-001
 ; CURRENT APPLICATION NUMBER: US/09/726,806
 ; CURRENT FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: 60/168,135
 ; PRIOR FILING DATE: 1999-11-30
 ; NUMBER OF SEQ ID NOS: 6283
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4625
 ; LENGTH: 3309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-726-806-4625

Alignment Scores:
 Pred. No.: 4.4e-159 Length: 3309
 Score: 2195.50 Matches: 416
 Percent Similarity: 84.97% Conservative: 8
 Best Local Similarity: 83.37% Mismatches: 10
 Query Match: 84.83% Indels: 65
 DB: 29 Gaps: 4

US-09-902-772-4 (1-478) x US-09-726-806-4625 (1-3309)
 OY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 DB 128 ATGGCCACACTATTAAAGGAAGCCCTTATCAGTTGTGATGAGGACCACTCGCTTTGAG 187

```

QY 21 CysAlaTyrGlySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB 188 TGTGCTACGGAAGCCACACCTGGCTTAACAGAGATGACCCGCTCTCTCCAGGAC 247
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
DB 248 TATGAGACAGACTTCGAAGATGAGCCACGCGCTCCCTACGAGATGGCTGTCTCAACCC 307
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAAGCCAGGGTCCACCAAAATGGAATGTAACCTTACCCAGGTGATGCTCAAGAAC 367
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB 368 TCTCCGTATGATGCACTGCTGGCCAAAGCCGGAAGATGGTGGCCAGCCAGACCCGTT 427
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
DB 428 GGGATGAACTACGAGCTACATGAGAGAGACACATGCCACCCCAACATGACACG 487
QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
DB 488 AACGAGCGCAGAGTTATCTGCGCAGATCTTACCGTATGAGTACAGACATGTGCGG 547
QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB 548 CAGTGGCTGGAGTGGGGCGTGAAGAATATGGCTTCCAGACCTCAACATCTTGTATTC 607
QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
DB 608 CAGAACATCGATGGGAAGCAACTGTGCAAGATCAACCAAGACGACCTCCAGAGCTCAC 667
QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
DB 668 CCCAGCTACACGCCGACATCTTCTCTCACTCTCCATCTCCAGAGAGACTCTCTT 727
QY 201 ProHISLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
DB 728 CCACATTTGACTTCAGATGATGTGATTAAGCTTACAAACCTCTCCAGCGTTAAATGAT 787
QY 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
DB 788 GCTAGAAACACA----- 799
QY 241 ThrGlnArgIleThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrp 260
DB 800 -----GATTACCATATGAGCCGCCAGGAATCAGCTGG 835
QY 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrVal 279
DB 836 ACCGGTCACGGCCACCCACGCGCCAGTCGAAAGCTGCTCAACCATCTCTTCCACAGTG 895
QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
DB 896 CCCAAAACGTGAACAGCAGCTCTCTCAGTTAGATCTTATCAGATCTTGGACCAACAGT 955
QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeu 319
DB 956 AGCCGCTTGTCAATCCAGGACAGTGGCCACATCCACTTTGGCAGTTCCTCCGAGCTC 1015
QY 320 LeuSerAspSerSerAsnSerAsnGlyLeuThrTrpGluGlnThrAsnGlyLysIlePheLys 339
DB 1016 CTGTGGACAGCTTCAACTCCAGCTGATATCCTGGGAAGGCCAAGGGGAGTTCAAG 1075
QY 340 MetThrAspProAspGluValAlaIleArgArgTrpGlyLysArgLysSerLysProAsnMet 359
DB 1076 ATACAGGATCCGACAGAGTGGCCCGCGCTGGGAGAGCGGAAGCAACCAACATG 1135
QY 360 AsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
DB 1136 AACTTAGATTAAGTCAAGCCCGCCCTCCCTACTATATGCAAGAACATCAACAGACCAAG 1195
QY 380 ValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399

```

```

DB 1196 GTC----- 1198
QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
DB 1199 ---CATGGGAAGCGCTACGGCTCAAGTGTGACTTCACAGGAGATGCCAGGCCCTCAG 1255
QY 420 Pro----- 420
DB 1256 CCCACACCCCGGAGTCACTGTACAGTACCCCTCAGACCTCCGTAATGAGGCTCC 1315
QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
DB 1316 TATCAGCGCCACCCACAGAAAGATGAACTTGTGGCGCCCAACCTCCAGCCCTCCCGTG 1375
QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIle 459
DB 1376 ACATCTTCCAGTTTGTGCTGCCCAACCCCTACGTAATTCACCAACTGGGGGTATA 1435
QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
DB 1436 TACCCCAACACTAGGCTCCACACGCAATATGCTTCTGATCTGGGACTACTAC 1492

RESULT 7
US-09-726-811-4920
; Sequence 4920, Application US/09726811
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; APPLICANT: Welch, Nadine S.
; APPLICANT: Wen, Danyl
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2027-001
; CURRENT APPLICATION NUMBER: US/09/726,811
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/168,136
; NUMBER OF SEQ ID NOS: 5515
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4920
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-811-4920

Alignment Scores:
Pred. No.: 4,4e-159 Length: 3309
Score: 2195.50 Matches: 416
Percent Similarity: 84.97% Conservative: 8
Best Local Similarity: 83.37% Mismatches: 10
Query Match: 84.83% Indels: 65
DB: 29 Gaps: 4

US-09-902-772-4 (1-478) x US-09-726-811-4920 (1-3309)
QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGlnAspGlnSerLeuPheGln 20
DB 128 ATGGCCAGCACTATTAAGGAAGCTTATCATGTTGTAGTAGAGACCAAGTCTGTTGAG 187
QY 21 CysAlaTyrGlySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB 188 TGTGCTACGGAAGCCACACCTGGCTTAACAGAGATGACCCGCTCTCTCCAGGAC 247
QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
DB 248 TATGAGACAGACTTCGAAGATGAGCCACGCGCTCCCTACGAGATGGCTGTCTCAACCC 307
QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 308 CCAAGCCAGGGTCCACCAAAATGGAATGTAACCTTACCCAGGTGATGCTCAAGAAC 367
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100

```


Db	368	TCCTCGTGAATGACAGTGTGGCCAAAGGCGGGAAGATGTTGGGAGCCACAGACCGTT	427
OY	101	GIYMeIAaNTyGlySerTyMeIuGIuYhShISleProProProAsnMeThrThr	120
Db	428	GGGATGAACCTACGGGAGCTACTGTGAGGAGGACGACCCGCCCAAACTGTGACAG	487
OY	121	ASngIuArGvValIleValProIAAsPProThrLeuTPSerThAspISValArg	140
Db	488	AACGAGCCCAAGTATTGTGTGCCACAGATCCTACGCTATGGAGTACAGACATGTGGG	547
OY	141	GIuTPLeuGIuTPAlaValIysGIuTYrGILeuProAsPValAspIleLeuLeuPhe	160
Db	548	CAGTGGCTGGAGTGGCGGCTGAAMAATAATGGCCCTCCAGACGTCAACATCTTGTATTC	607
OY	161	GIuAsnIleAsPGLyGIuLeuGlyuLeuYsIysMeThrIlyAsAsPPhelIaArgLeuThr	180
Db	608	CAGAACATCATGGGAAGGAACTGTGCAGAGATGACCAAGGAGAGACTTCCAGAGGCTCACG	667
OY	181	ProSerTyAsnAlaAspIleLeuLeuSerHisIleuHisTyLeuArGIuThrProLeu	200
Db	668	CCACACTCAAGAGCGGACATCCTGTCTCACATCTCCACTACCTCAGAGAGACTCCTGT	727
OY	201	ProHisLeuThrSerAspAspValAspIlySalAlaLeuGIuAsnSerProArgLeuMetHis	220
Db	728	CCACATTTGACTCTGATGATGTTATGAAGCTTACAAACTCCACAGGTTAATGCAT	787
OY	221	AlaArGAsnThrGIyGIyAlaThrPheIlePheProAsnThrSerValTyProGIuAla	240
Db	788	GCTAAACACAC-----	799
OY	241	ThrGIuArGIleThrThrArgProAsPLeuProTYrGIuGIuAlaIaArgArgSerAlaTrp	260
Db	800	-----GATTTACCATGTGAGCCCCCGGAGAGATCACCCGG	835
OY	261	ThrSerHisSerHisAsPProThr--GIuSerIlySalAThGIuProSerSerThrVal	279
Db	836	AACGGTCAACGGCCACCCACGCGCCAGTGCAGAAAGTGCCTCAACATCTCTCCACAGTg	895
OY	280	ProIlyThrGIuAsPGLaArgProGIuLeuAsPProTYrGIuIleLeuGIyProThrSer	299
Db	896	CCCAAACTGGAAGACACAGCTCCACGTACGTACTCTTCAGATTCTTGACCAACAAGT	955
OY	300	SerArgIleuAlaAsnProGIySerGIuIleGIuIleuATPGLInPheLeuLeuLeu	319
Db	956	AGCGGCTTGCAAACTCCAGGAGTGGCCGAGATCAGCTTTGGCACTTCTCTCGAAGCTC	1011
OY	320	LeuSerAsPserSerAsnSerAsnCySleThrTrpGIuGIyThrAsnGIyGIuPheIys	339
Db	1016	CTGTGCGACACTCCAACTCCAGCTGCATCACTCGGAAGGACCAAGGGGAGTTCAAG	1075
OY	340	MeThrAsPProAsPGLuValAlaArgArGIrPGLyGIuArgIysSerIysProAsnMet	359
Db	1076	ATGACGGGATCCCGAGAGGTGCCCCGGCGTGGGAGGAGCGMAAGCAAAACCCAAACTG	1135
OY	360	AsnTyAsPlySLeuSerArGIAlaLeuArGIrTyTYrTYrAsPlyAsnIleMeThrIys	379
Db	1136	AACACAGATTAAGCTCAGCGGCGCCCTCCGTTACTACTATGACAGAAACATCATGACCAAG	1195
OY	380	ValHisProProGIuSerSerMeTrIlyTyTYrProSerAsPLeuProTYrMetSerSer	399
Db	1196	GTC-----	1198
OY	400	TyrHisGIySArGIYrAlaTYrIlySPhaSPheHisGIyIleAlaGIuAlaLeuGIu	419
Db	1199	---CATGGGAAGCGCTACGCTTACAAGTTGCACTTCCAGGGGATGCCCCAGGCGCTCCAG	1255
OY	420	Pro-----	420
Db	1256	CCCCACCCCGGAGTCACTCTGTACAAAGTACCCTCAGACCTCCGTACATGGGCTGCC	1315
OY	421	---HisAlaHisProGIuIlyMetAsnPhelAlaIaProHisProProAlaLeuProVal	439
Db	1316	TATACAGCCCAACACAGAAATGAACTTTGTGGGCCCAACCCCTCCAGCCCTCCCGGTG	1375

QY	440	ThrsrserseerPhePheaaiaaiaProaSnProTyrTrpAsnSerProThgIylyle	455
Db	1376	ACATCTCCAGATTTTTTGTCTGCCCAAAACCCTACTGGAAATTCACCAATGGGGGTATA	1436
QY	460	TyPProaSnThirArgLeuProaiaaiaIaHsMeIProSerHisLeuGIyThrTyTyf	478
Db	1436	TACCCCAACACTAGCGTCCCCACACACCATATGCCCTTTCATCTGGGCACTTACTAC	1492
RESULT 8			
US-09-948-941-233			
: Sequence 233, Application US/09948941			
: GENERAL INFORMATION:			
: APPLICANT: VENTER, J. Craig et al.			
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
: TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF			
: FILE REFERENCE: CL000788			
: CURRENT APPLICATION NUMBER: US/09/948,941			
: CURRENT FILING DATE: 2001-09-10			
: PRIOR APPLICATION NUMBER: 60/231,328			
: PRIOR FILING DATE:			
: NUMBER OF SEQ ID NOS: 12618			
: SOFTWARE: FastSeq for Windows Version 4.0			
: SEQ ID NO 233			
: LENGTH: 1389			
: TYPE: DNA			
: ORGANISM: Human			
US-09-948-941-233			
Alignment Scores:			
Pred. No.: 2,22e-158 Length: 1389			
Score: 2180.50 Matches: 413			
Percent Similarity: 84.74% Conservative: 9			
Best Local Similarity: 82.93% Mismatches: 11			
Query Match: 84.23% Indels: 65			
DB: 35 Gaps: 4			
US-09-902-772-4 (1-478) x US-09-948-941-233 (1-1389)			
QY	2	AlaserthrIlelySGlualaleuSerValaSeGIuaSPGInSerLeuPheGIuCyS	21
Db	25	GCACCTCATATTCAGAGGAAGGCTTATGATGTGAGAGAGAGACCACTGTTGTGAGTGT	84
QY	22	AlatyGIySerProHIsleuAlaIySTrGIuMetThrAlaSerSerSerSeGIuTyR	41
Db	85	GCCACGGAACGCCACACCTGGCTAAGAGAGATGACCGCTCCTCCACGAGCAT	144
QY	42	GIyInThrSerIyMetSerProArGValProGIInaSPTrpLeuSerGIInProPro	61
Db	145	GGACAGACTTCCAAGATGAGCCACGCGCTCCACAGAGATGGCTGTCTCAACCCCA	204
QY	62	AlaArgValthrIlelyMetGIuCySaSProaSnGInValaSnGIySerArGAsnSer	81
Db	205	GCCAGGGTCCCATCAAAATGGAATGTACCTACCAAGGTGAATGGCTCAAGGAAGCTT	264
QY	82	ProaSPaSPCySerValAlaIySGIyGIyMetValaSerSerSerSerPaSnValGIy	101
Db	265	CCGTGATCAATGCGAGTGTGGCCAAAGCGGGAAGATGTGTGGCAGCCGACACCGTTGGG	324
QY	102	MetAsnTyGIySerTyMetGIuGIuIySHIsIleProProProaSnMetThrThrAsn	121
Db	325	ATGAACACGCGAGCTACATGTGAGAGGAGACACATGCCACCCCAACATGACACGAAAC	384
QY	122	GIuArgArgValIleValProaIaSPProIleuTrpSerThrAspHisValArgGIn	144
Db	385	GAGCGCAGATTTTCCTGGCAGAGAGTCTTACGCTATGAGACAGACCATGTGGCGAG	444
QY	142	TrpLeuGIuTrpAlaValaIySGIuTyrGIyLeuProaSPValaSPIleLeuLeuPheGIn	161
Db	445	TGGCTGAGTGGCGGCTGAAGAATATGGCTTCACAGACGTCAACATCTTGTATTCCAG	504
QY	162	AsnIleAspGIyGIyGIuLeuCySIySMeThrTySaSPaSPheGInaIyLeuThrPro	181


```

Db 505 AACATCATGGGAAGAACTGTGCAAGATGACCAAGACGACTTCAGAGGCTCACCCCC
Qy 182 SerTyraSnlaAspIleLeuSerHisIleuHisTyrlleuArgIuThProLeuPro 201
Db 565 AGCTACACAGCCGCACTCTCTCTCATCTCCATCTCCAGATCCAGAGAGCTCTTCCA
Qy 202 HisLeuThrSerAspAspValAspLysAlaIleuGlnInsSerProAgluMethisAla 221
Db 625 CATTTGACTTCAGATGATGTTGATTAAGCCCTTACAAAACCTCCACGGTTATATCATGCT 684
Qy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 685 AGAAACACACA-----
Qy 242 GlnArgIleThrThrArgProAspLeuProGlyGluGlnAlaArgAspSerAlaTyrThr 261
Db 694 -----GATTTACCATATGAGCCGCCCAAGAAATCAGCTTGAGCC
Qy 262 SerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 733 GGTACAGGCGCCAGCCAGCCAGCTCGAAGAACCTCTCAGACATCTCTCCACAGTGTCCC 792
Qy 281 LysThrGluAspGlnArgProGlnLeuAspProGlyGlnIleuGlyProThrSerSer 300
Db 793 AAAACTGAAGACAGCGCTCTCAGTATGATCTTATGATTCAGATTCTTGAGACCAAGATAGC 852
Qy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGlnLeu 320
Db 853 CGCCTTGCAATTCAGGAGGAGGCGCCAGATCCAGCTTTGGCGATTCTCTCGAGCTCTCG 912
Qy 321 SerAspSerSerAsnSerAsnGlyIleThrTrpGluGlnIleThrAsnGlyLysPheLysMet 340
Db 913 TCGGACAGCTCCACATCTCAGTGCATCAGCCGGGAAGACCAAGCGGGAGTTCAAGATG 972
Qy 341 ThrAspProAspGluValAlaIleArgTyrGlyLysLysSerLysProAsnMetAsn 360
Db 973 ACGGATCCCGACGAGGAGGCGCGCGCTGGGAGAGAGCGGAAGCAAAACCAATCATGAAAC 1032
Qy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db 1033 TACATTAAGCTACAGCCGCGCTCTCGTACTACTATGACAAAGAACATCATGACCAAGGTC 1092
Qy 381 HisProGluSerSerMetLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1092 -----
Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1093 CATGGGAAGCGCTACGCTACAAAGTTGCACTTCCACGGGATCGCCAGGCCCTCCAGGCC 1152
Qy 420 -----
Db 1153 CACCCCGGAGTCATCTGTACAGTACCCCTCAGACCTCCGTCATATGGGCTCTAT 1212
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1213 CAGCGCACCCACAGAGATGAATTTGTGGCGCCACACCTCCAGGCCCTCCCGGTACACA 1272
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTyrAspSerProThrGlyLysIleTyr 460
Db 1273 TCTTCCAGTTTGTGTGGCCCAAAACCATACGTGAATTCACCACTGGGGGTATATAC 1332
Qy 461 ProAspThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1333 CCCAACACTAGGCTCCCAACAGCATATGCTTCTCATCTGGGCGACTTACTAC 1386

```

```

; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044PC
; CURRENT APPLICATION NUMBER: PCT/US02/23913
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-23913-98

Alignment Scores:
Pred. No.: 6,02e-158 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 1 Gaps: 4

US-09-902-772-4 (1-478) x PCT-US02-23913-98 (1-3166)
Qy 2 AlaSerThrIleLysGluAlaLeuSerValLysGluAspGlnSerLeuPheGluLys 21
Db 281 GCACCTATATCAAGAGAGCCCTTATCAGTTGTGAGTAGGACCACTGTGTGTGAGTGT 340
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlyTyr 41
Db 341 GCCTACGGAAGCCACACCTGTGTAAAGACAGATGACCGGCTCTCTCCAGGACTAT 400
Qy 42 GlyIleThrSerLysMetSerProArgValProGlnGlnAspTyrPheLeuSerGlnPro 61
Db 401 GACAGACTTCCAGATGAGAGCCAGCGCTCCCTCAGCAGATGAGTGTGCTCAACCCCA 460
Qy 62 AlaArgValThrIleLysMetGluLysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 461 GCCAGGCTCACCATCAAAATGGAATGTAAACCTTACGCCAGGTGAATGCTCAAGAACTCT 520
Qy 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 521 CTTATGATATGCACTGTGTGCCAAGAGCGGGAAGTGTGGCCAGCCACAGACCCGTTGGG 580
Qy 102 MetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThrAsn 121
Db 581 ATGAATACAGGACACTGATGAGAGAGACATGACACACCCCAAAATGACACACAGMAC 640
Qy 122 GlnArgArgValIleValProAlaAspProThrIleuThrPheSerThrAspHisValArgGln 141
Db 641 GAGGACAGAGTATACGTGCCAGCAGATCTTACGCTATGATGATGACACCAATGTCCGAG 700
Qy 142 TrpLeuAluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161
Db 701 TGGCTGAGTGGGCGGTGAAGAAATATAGGCTTCAACAGTCACATCTGTATTTCAG 760
Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATGATGGAAGAGAACTGTGCAGATGACCAAGAGACACTCCAGAGGCTCACCCCC 820
Qy 182 SerTyraSnlaAspIleLeuSerHisIleuHisTyrlleuArgIuThProLeuPro 201
Db 821 AGCTACACAGCCGCACTCTCTCTCATCTCCATCTCCAGATCCAGAGAGCTCTTCCA 880
Qy 202 HisLeuThrSerAspAspValAspLysAlaIleuGlnInsSerProAgluMethisAla 221

```

RESULT 9
 PCT-US02-23913-98
 : Sequence 98, Application PC/TUS0223913
 : GENERAL INFORMATION:

: APPLICANT: Millennium Pharmaceuticals, Inc. et al.
 : TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 : TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

```

Db      881  CATTGACTCAGATGATGTTGATAAGCCTTACAAAACCTCCACGGTATGACAGCT 940
Qy      222  ArgSnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db      941  AGAACAACA----- 949
Qy      242  GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaThrPhe 261
Db      950  -----GATTACCATATAGAGCCGCCCCAGAGATGACCTGGAGCC 988
Qy      262  SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      989  GGTACAGCGCCACCCACAGCCCGCCAGCAAGCTCTCAACCATCTCTCCACAGATGCC 1048
Qy      281  LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      1049  AAAACCTAAGACAGCGCTCCATGATCCATTATGATGATTTGGACCAACAAGTAGC 1108
Qy      301  ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheGlnPheLeuGluLeuLeu 320
Db      1109  CGCTTCGAATCCAGGCGAGTGGCCAGATCCAGCTTGGCAGTTCCTCCGTGAGACTCTCG 1168
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTyrGluGlyThrAsnGlyGluPheLysMet 340
Db      1169  TCGGACAGCTCCACATCCAGCTGATCAGCTGGAAGGACCAACGAGGAGTTCAGATG 1228
Qy      341  ThrAspProAspGluValAlaArgArgTyrGlyLysThrGlyLysSerLysProAsnMetAsn 360
Db      1229  AGCGATCCCGACGAGGGGGCGGCGTGGGAGAGCGGAAGCAACCCCAACATGAGAC 1288
Qy      361  TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db      1289  TAGCATAGCTCAGCCCGCGCTCGTTACTACTATGACAAGAACATCATGACCAAGTGC 1348
Qy      381  HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1348  ----- 1348
Qy      401  HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1349  CATGGAGAGCGCTACGCTTCAAGATTGACATCCAGCGGATCGCCCGCCCTCCAGCCC 1408
Qy      420  ----- 420
Db      1409  CACCCCGGAGATCTCTGTACAAGTACCCCTCAGACCTCCGCTACATGGCGCTCTAT 1468
Qy      421  HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1469  CAGCCCGACCCACAGAGATGAATTTGTGGCGCCACCCCTCCAGCCCTCCCGGTACA 1528
Qy      441  SerSerSerPhePheAlaAlaProAsnProTyrTyrAsnSerProThrGlyGlyIleTyr 460
Db      1529  TCTTCCAGTTTTTTGTGTCGCCCAACCCATACGGAATTCACCAACTGGGGGTATATAC 1588
Qy      461  ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db      1589  CCCAACACTAGGCTCCCCACAGCATATAGCTTCATCTGGGCACTTACTAC 1642

```

```

; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-356-26

Alignment Scores:
Pred. No.: 6,02e-158 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: Gaps: 4

US-09-902-772-4 (1-478) x US-09-784-356-26 (1-3166)
Qy      2  AlasThrIleLysGluAlaLeuSerValValSerLysGlnSerLeuPheGluCys 21
Db      281  GCACCTCATATCAAGGAAGCCTTATCAGTTGTGATGAGACCCAGTGTGTTGAGTGT 340
Qy      22  AlatyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41
Db      341  GCCTAGGAAAGCCACCTGGCTGAGACAGATGACCGGCTCTCTCCAGCGACTAT 400
Qy      42  GlyLysThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGlnProPro 61
Db      401  GGACAGACTTCCAAAGATGAGAGCCCGGCTCCCTCAGCAGATGGCTGTCAACCCCA 460
Qy      62  AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerThrArgAsnSer 81
Db      461  GCCAGGGTCAACCAAAATGGAATGTAAACCTAGCCAGCTGAATGGCTCAAGCAACTCT 520
Qy      82  ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db      521  CCGTATCAATGCACTGTGGCCAAAGCGGGAAGATGGTGGCAGACCCAGACCGGTGGG 580
Qy      102  MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
Db      581  ATGAACCTACGCGAGCTCATGTGAGAGAGAACACATGCCACCCCAACATGACACGAGAC 640
Qy      122  GlnArgArgValIleValProAlaAspProThrLeuThrPheThrAspHisValArgGln 141
Db      641  GAGCGAGAGTATATCGCCAGCAGATCTTACGCTATGAGATACAGACCATGTGCGGACG 700
Qy      142  TyrLeuGluTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161
Db      701  TGGCTGAGTGGGGGGGGAAGAATATGAGCTTCCAGACGTCAACATCTTGTATTCCAG 760
Qy      162  AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      761  AACATCGATGGGAAGAACTGTGCAGATGACCAAGAGACACTCCAGAGGCTCACCCCC 820
Qy      821  AGCTACAAAGCCGACATCTTCTCTCACTCCACTCACTCCAGAGAGATCTCTCTCCA 880
Db      182  SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
Qy      881  CATTGACTTCAAGTATGTTGATTAAGCTTACAAAACTCTCCAGCGTTAAAGCATGCT 940
Db      941  AGAACAACA----- 949
Qy      222  ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db      242  GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaThrPhe 261
Qy      950  -----GATTACCATATGAGCCGCCCCAGAGATGACCTGGAGCC 988

```



```

Db 1109 CGCCTTCGAATCCAGCAGTGGCCAGATCCAGCTTGGCGAGTCTCCCTGAGCTCCTG 1168
Qy 321 SerAspSerSerAsnSerAsnCysIleThrTrpIleuGlyThrAsnGlyIleuPhenylMet 340
Db 1169 TCGGACAGCTCCAACTCCAGCTGCATCAGCTGGGAGAGCACCAGGGGAGTTCAAGATG 1228
Qy 341 ThrAspProAspGluValAlaIleArgArgTrpGlyIleuArgLysSerLysProAsnMetAsn 360
Db 1229 ACGGATCCCGACGAGGTGGCGCCGCTGGGGAAGACGGAGAACCAACCAATGAC 1288
Qy 361 TyrAspLysLeuSerArgAlaLeuArgLysTrpTrpLysAspLysAsnIleMetThrLysVal 380
Db 1289 TACCATAGCTCACCCCGCCCTCCGTACTATATGACAAGAACATCATGACCAAGTTC 1348
Qy 381 HisProGluSerSerMetTyrLysTyrProSerLysPheProTyrMetSerSerTyr 400
Db 1348 ----- 1348
Qy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1349 CATGGAGAGCGCTACGCTACAAAGTTGAGCTTCACAGGGATCGCCAGGCCCTCCAGCCC 1408
Qy 420 ----- 420
Db 1409 CACCCCGGAGTATCTGTACAGTACCCCTCAGACCTCCGCTACATGGCTCTAT 1468
Qy 421 HisAlaHisProGlnLysMetAsnPheValAlaIleProHisProProAlaLeuProValThr 440
Db 1469 CAGCCGCCACCCACGAGAGTAATCTTGTGGCGCCCAACCTCCAGCCCTCCCGTGACA 1528
Qy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIleTyr 460
Db 1529 TCTTCAGTATTTTGTGTCGCCCAACCAACCATGATGCAATTCACCACTGGGGGTATATAC 1588
Qy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1589 CCCAACACTAGCTCCGCCACCAACCATATGCTTCATCTGAGGCGACTTACTAC 1642

RESULT 12
US-10-205-823-98
: Sequence 98, Application US/10205823
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Monahan, John E.
: APPLICANT: Endege, Wilson O.
: APPLICANT: Gannavarapu, Manjula
: APPLICANT: Gorbacheva, Bella
: APPLICANT: Hoersch, Sebastian
: APPLICANT: Kamatkar, Shubhangi
: APPLICANT: Monsey, Angela M.
: APPLICANT: Gialt, Karen
: APPLICANT: Zhao, Xumei
: APPLICANT: Anderson, Dustin
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
: TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: FILE REFERENCE: MRI-044
: CURRENT APPLICATION NUMBER: US/10/205, 823
: PRIOR APPLICATION NUMBER: 2002-07-25
: PRIOR FILING DATE: 2001-07-25
: PRIOR APPLICATION NUMBER: 60/314,356
: PRIOR FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/325,020
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: 60/341,746
: PRIOR FILING DATE: 2001-12-12
: PRIOR APPLICATION NUMBER: 60/362,158
: PRIOR FILING DATE: 2002-03-05
: NUMBER OF SEQ ID NOS: 455
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 98

```

```

: LENGTH: 3166
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-205-823-98

Alignment Scores:
Pred. No.: 6,02e-158 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 42 Gaps: 4

US-09-902-772-4 (1-478) x US-10-205-823-98 (1-3166)

Qy 2 AlaSerThrIleLysGluAlaLeuSerValValSerLysLysAspGlnSerLeuPheGluCys 21
Db 281 GCAGCTCATATCAAGGAGGCTTATATGAGTGTGAGTGAAGACCAAGTGTGAGTGTG 340
Qy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
Db 341 GCCTACGAGAACGCCACACCTGGCTAAGACAGATGACCGCTCTCTCCAGGAGATAT 400
Qy 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnProPro 61
Db 401 GGACAGACTTCCAAAGATGAGCCCGCCCTCCAGAGATGGCTGTCAACCCCA 460
Qy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 461 GCCAGGCTCACCAATATGAAATGATATACCTTACCCAGGTGAATGGCTCAAGGACTCT 520
Qy 82 ProAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
Db 521 CCTGATGAATGACAGTGTGGCCAAAGCGGGGAAGATGGTGGGACGCCAGACCGTTGG 580
Qy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
Db 581 ATGAACATACGCGACGCTCATGAGAGAACACATGACCCCAACATGACACCAAC 640
Qy 122 GluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArgGln 141
Db 641 GAGCGCAGATTTATGTCGACGAGATCTTACGCTTACGAGTACAGCCATGAGCGGACG 700
Qy 142 TrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 701 TGGCTGAGTGGGGGCGTGAAGAATATGAGCTTCCAGACTCAACATCTGTTATTCAG 760
Qy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 761 AACATCGATGGGAAGAACTGTGCAGATGACCAAGACGACATCTCCAGAGCGTCACCCC 820
Qy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
Db 821 AGCTACACGCCCGACATCTTCTCTCAGCTCTCAGCTCTCAGAGAGATCTCTTCCA 880
Qy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 881 CATTTGACTTCAGATGATGTTGATTAAGCTTCAAAAACCTCCAGCGTTAATGACATGT 940
Qy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 941 AGAAACACA----- 949
Qy 242 GluArgGlnThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThr 261
Db 950 -----GATTTACCATATGAGGCCCGCCAGGAGATGAGCTGAGACC 968
Qy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 989 GGTACAGGCGCACCCAGCCCGCAGTGAAGAGCTGTCAACCATCTCTTCCACAGTGGCC 1048
Qy 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300

```

```

Db      1049 AAAAGTGAAGACGAGCTCTCAGTAGATCCCTTATCATGATCTTGGACCAACAAGTAGC 1108
Qy      301  ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglInPheLeuGluLeuLeu 320
Db      1109 CGCCTTGCAAAATCCAGAGGAGTGGCCAGATCCAGCTTGGCAGTTCTCTGGAGCTCTG 1168
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet 340
Db      1169 TCGGACAGCTCCAACTCCAGTCAGTCATCCCGGGAAGGACCAACGGGGAGTTCAAGATG 1228
Qy      341  ThrAspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsn 360
Db      1229 ACGGATCCCGACGAGGAGTGGCCCGGCTGGGAGAGCGGAAGCAACCAACCAATGTAAC 1288
Qy      361  TyrAspLysLeuSerArgAlaLeuArgTrpTyrTrpAspLysAsnIleMetThrLysVal 380
Db      1289 TACATATAGCTACACCGGCGCTCGTACTACTATGACAGAAACATCATGACCAAGGTC 1348
Qy      381  HisProProGluSerSerMetLysLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db      1348 ----- 1348
Qy      401  HisGlyLysArgTrpAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1349 CATGGGAAGCGCTACGGCTTACAAAGTTCAGCTTCCACGGGATCGCCAGGCCCTCCAGCCC 1408
Qy      420 ----- 420
Db      1409 CACCCCGGAGTCATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGGCTCTAT 1468
Qy      421  HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1469 CAGCCACCCACGACGAGATGAATCTTGTGGCCGCCACCCCTCAGGCCCTCCCGGTACA 1528
Qy      441  SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProHisGlyLysIleTyr 460
Db      1529 TCTTCCAGTTTCTTGTGGCCCAACCCATCTGGAATTCACCACTGGGGGGTATATAC 1588
Qy      461  ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db      1589 CCCAACACTAGCTCCCCACGAGCATTATGCTTCTCATCTGGGCACTTACTAC 1642

RESULT 13
US-60-207-355
; Sequence 355, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: C0001321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 355
; LENGTH: 2172
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-355

Alignment Scores:
Pred. No.: 1,52e-145 Length: 2172
Score: 2017.50 Matches: 384
Percent Similarity: 79.12% Conservative: 10
Best Local Similarity: 77.11% Mismatches: 7
Query Match: 77.96% Indels: 97
DB: 80 Gaps: 3

US-09-902-772-4 (1-478) x US-60-360-207-355 (1-2172)
Qy      1  MetaLysThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db      107  ATGGCCACGACTATTAAAGAGGCTTCTCAGTTGTGACGAGGACGATCACTATTGTAG 166
Qy      21  CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlu 40

```

```

Db      167  TGTGCTACGGAAGCCACACCTGGCTAAGACAGATGACCGCATCTCTTCCAGTGAC 226
Qy      41  TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
Db      227  TATGGCCAGATCCCAAGATGAGTCCAGATCTCCCTCAGCGAGCTGCTTCTTCAAGCC 286
Qy      61  ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
Db      287  CCAGCCAGGCTCACCATCATAGATGAGTGCACACCTTGTACAGTGAATGTTCCAGGAC 346
Qy      81  SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
Db      347  TCACCTGATGATGTCAGATGTAACAAGGTGGAAAGATGTTGGGACAGCCGATACGTG 406
Qy      101  GlyMetAsnTyrGlySerTyrMetGluGluHisIleProProProAsnMetThr 120
Db      407  GGGATGAGCTACGGCAGCTACATGAGAGAGATGTGCCCTCCCAATATACACACA 466
Qy      121  AsnGluArgArgValIleValProAlaAspProThrLeuThrPsrThrAspHisValArg 140
Db      467  AATAGCCGACAGATGATCTCTCTGACAGATCTCTGTGGACACACAGACATGTCCGA 526
Qy      141  GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLysPhe 160
Db      527  CAGTGGCTGGAGTGGGGGTGAAGAATATGGCTCTCCGATGTGAGAGCTTACTATT 586
Qy      161  GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
Db      587  CAGAAATATCGATGGGAGGAGCTGTGCAGATGTCAAAAGATGACTTCCAGCGGCTCAGC 646
Qy      181  ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
Db      647  CCGAGCTACAAATGCCGACATTTCTTCTACATCTCCACATCTCCAGAGACATCCCTT 706
Qy      201  ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLysMetHis 220
Db      707  CCACATCTGACTTCCGATGAGCTGTATAGCTTTACAAAACCTCCACAGGTTAATGCAT 766
Qy      221  AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
Db      767  GCCAGAAACACAGGGGGTGCAGCTTATTTTCCAAATATTCTAGATATATCCGAGAGCT 826
Qy      241  ThrGlnArgIleThrTrpArgProAspLeuProTyrGluGlnAlaArgSerAlaTyr 260
Db      827  ACGCAAAGAATTACAACTAGG----- 847
Qy      261  ThrSerHisSerHisProHisGlnSerLysAlaThrGlnProSerSerThrValPro 280
Db      847 ----- 847
Qy      281  LysThrLysAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProHisSerSer 300
Db      847 ----- 847
Qy      301  ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPglInPheLeuGluLeuLeu 320
Db      848 ----- 848
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet 340
Db      896  TCGAGACAGCTCCAACTCCAGTCAGTCACCTGGGAAGGACCAACGGGGAGTTCAAGATG 955
Qy      341  ThrAspProAspGluValAlaIleArgTrpGlyGluArgLysSerLysProAsnMetAsn 360
Db      956  ACAAGACCCGAGCAGGAGTGGCTCGGCTGGGGGAGAGGAAGCAACCAACATGTAAC 1015
Qy      361  TyrAspLysLeuSerArgAlaLeuArgTrpTyrTrpAspLysAsnIleMetThrLysVal 380
Db      1016  TATGACAGCTCAGCCGCGCTCTGCTACTACTACACAAAAACATCATGACCAAGGTG 1075
Qy      381  HisProProGluSerSerMetLysLysTyrProSerAspLeuProTyrMetSerSerTyr 400

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:30:39 ; Search time 1092.35 Seconds

(Without alignments)
3233.216 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKREALSVSEDSLEF.....IYPNRLPAHMHSHLGTYY 478

Scoring table:

BIOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Delop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 8407786 segs, 3694357880 residues

Total number of hits satisfying chosen parameters: 16815572

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL-frame+2pn.model -DEV-rlh
-Q=/cgn2_1/USPto.spool/US09902772/runat_23072003_093658_14943/app_query.fasta.1.1294
-DB-pending-Patents_NA.New -QFMT-fastap -SUFFIX-trnp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITs-bits -START=1 -END=1 -MATRIX-biosum62 -FRANS-human40.cdl
-LIST=45 -DOCALLIG=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=ptc -NOM-ext -HEAPSIZE=500 -MINTER=0 -MAXLEN=200000000
-USER=US09902772.GCN_1.1.325.grunat.23072003_093658_14943 -NCPu=6 -ICPU=3
-NO_MMAP -LARGEIOBERT -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database :

```

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	2324.5	89.8	1930	15	US-10-170-235-28132	Sequence 28132, A
2	2324.5	89.8	1930	18	US-60-452-680-11396	Sequence 11396, A
3	2195.5	84.8	3294	15	US-10-170-235-28519	Sequence 28519, A
4	2195.5	84.8	3294	18	US-60-452-680-11397	Sequence 11397, A
5	2180.5	84.3	1389	12	PCT-US02-04915-41	Sequence 41, Appl
6	2180.5	84.3	1389	14	US-10-211-4652-41	Sequence 41, Appl
7	2180.5	84.3	1389	15	US-10-087-192-1137	Sequence 1127, Ap
8	2180.5	84.3	1389	14	US-10-126-052A-330	Sequence 330, App
9	2083.5	80.5	1297	15	US-10-087-192-1124	Sequence 1124, Ap
10	2017.5	78.0	2172	14	US-10-144-771-355	Sequence 355, App
11	1482	57.3	1359	2	PCT-US02-41414-1210	Sequence 1209, Ap
12	1482	57.3	1359	2	PCT-US02-41414-1209	Sequence 710, App
13	1482	57.3	2957	14	US-10-342-887-710	Sequence 1207, Ap
14	1474	57.0	1359	2	PCT-US02-41414-1207	Sequence 1206, Ap
15	1474	57.0	1729	2	PCT-US02-41414-1206	Sequence 10845, A
16	1474	57.0	1729	14	US-10-144-771-10845	Sequence 34391, A
17	1468	56.7	2769	15	US-10-170-235-34391	Sequence 4688, Ap
18	1465	56.6	2172	11	US-09-949-016-4688	Sequence 1126, Ap
19	809.5	31.3	212231	15	US-10-087-192-1126	Sequence 1126, Ap
20	809.5	31.3	8059021	9	US-09-947-914-53	Sequence 53, Appl
21	802	31.0	69090	15	US-10-087-192-1123	Sequence 1123, Ap
22	801	31.0	567	14	US-10-203-138A-6819	Sequence 6819, Ap
23	716	27.7	473	14	US-10-203-138A-8888	Sequence 8889, Ap
24	702.5	27.1	57726	11	US-09-949-016-16430	Sequence 16430, A
25	702.5	27.1	73725	2	PCT-US02-41414-1208	Sequence 1208, Ap
26	685.5	26.5	72732	2	PCT-US02-41414-1205	Sequence 1205, Ap
27	654	25.3	533	10	US-09-513-999C-1340	Sequence 1340, Ap
28	654	25.3	533	12	US-09-513-999C-1340	Sequence 1340, Ap
29	503	19.4	420	14	US-10-203-138A-3751	Sequence 3751, Ap
30	503	19.4	454	14	US-10-203-138A-1690	Sequence 1690, Ap
31	498	19.2	318	10	US-09-513-999C-25950	Sequence 25950, A
32	498	19.2	318	12	US-09-513-999C-25950	Sequence 25950, A
33	481	18.6	1752	15	US-10-219-051B-13871	Sequence 13871, A
34	474.5	18.3	1876	15	US-10-170-235-24888	Sequence 24888, A
35	474.5	18.3	1876	18	US-60-452-680-2295	Sequence 2295, App
36	474.5	18.3	1901	15	US-10-219-051B-13873	Sequence 13873, A
37	469	18.1	1581	15	US-10-170-235-34170	Sequence 34170, A
38	469	18.1	1581	19	US-60-455-444-673	Sequence 673, App
39	469	18.1	1581	19	US-60-465-241-673	Sequence 673, App
40	464	17.9	1165	15	US-10-170-235-24404	Sequence 24404, A
41	464	17.9	1165	18	US-60-452-680-2296	Sequence 2296, App
42	448	17.3	443	14	US-10-144-771-21735	Sequence 21735, A
43	444	17.2	2422	15	US-10-170-235-34449	Sequence 34449, A
44	444	17.2	2422	15	US-60-452-680-11304	Sequence 11304, A
45	443.5	17.1	3769	18	US-10-170-235-28517	Sequence 28517, A

ALIGNMENTS

```

RESULT 1
US-10-170-235-28132
/ Sequence 28132, Application US/10170235
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig
/ TITLE OR INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU.
/ FILE REFERENCE: CL001380
/ TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
/ CURRENT APPLICATION NUMBER: US/10/170, 235
/ CURRENT FILING DATE: 2003-03-17
/ NUMBER OF SEQ ID NOS: 42514
/ SEQ ID NO 28132
/ LENGTH: 1930
/ TYPE: DNA
/ ORGANISM: HUMAN
/ US-10-170-235-28132

```

Alignment Scores:

Pred. No.:	5,34e-259	Length:	1930
Score:	2324.50	Matches:	436
Percent Similarity:	89.36%	Conservative:	9
Best Local Similarity:	87.55%	Mismatches:	12

Query Match: 89.82% Indels: 41
 DB: 15 Gaps: 3
 US-09-902-772-4 (1-478) x US-10-170-235-28132 (1-1930)

QY 2 AlaSerThrIleuysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
 DB 25 GCACCTCATATCAATGAGGAGGAGGCTTATGAGTGTGAGTGAGGACCAAGTGTGTTGAGAGCT 84
 QY 22 AlATyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
 DB 85 GCCTACGGAGACGCCACACCTGCTAGACAGATGACCGGCTCTCTCCAGCGACAT 144
 QY 42 GlyIlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
 DB 145 GGACAGACTTCCAAAGATGAGCCACGGCTCCCTCAGCAGATGGCTGTCAACCCCA 204
 QY 62 AlaArgValThrIleuysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 DB 205 GCCAGGGTCACCATCAAAATGGAATGTAACCTAGCCAGGTGAATGGCTCAAGGACCTCT 264
 QY 82 ProAspAspCysSerSerValAlaLysGlyLysMetValSerSerSerSerAspAsnValGly 101
 DB 265 CCTATGATGATGACATGTGTGGCCAAAGCGGGAAGTGGTGGCCACCCAGACACCGTTGG 324
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThrAsn 121
 DB 325 ATGAACACTACGCGACCTCATGTGAGAGACACATGCGACCCCAACATGACCAACAC 384
 QY 122 GluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArgGln 141
 DB 385 GAGGCGAGAGTTATCGTGCCAGCAGATCTCAGCTATGATGATACAGACCATGTGCGCAG 444
 QY 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAlaIleLeuLeuPheGln 161
 DB 445 TGGCTGGAGTGGGGGGAAGAATATGGCTTCCAGACGTCACATCTTTTTCAG 504
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
 DB 505 AACATCATGTGGAGGAGGAGTGTGCAGATGACCAAGACGACCTTCCAGAGCTCACCC 564
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
 DB 565 AGCTACACGCGGACATCTCTCTCAATCTCCACTACCTCAAGAGACCTCTTCCA 624
 QY 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
 DB 625 CATTTGACTTCAGATGATGTGATTAAGCCTTACAAACTCTCCAGCGTTAATCATGCT 684
 QY 222 ArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
 DB 685 AGAAACACAGGGGTGACAGCTTTATTTTCCAAATCACTTCAGATATCTCGAAGCTACG 744
 QY 242 GlnArgIleThrThrArgProAspLeuProTyrGlnLlnAlaArgArgSerAlaTrpThr 261
 DB 745 CAAGAATTTCACTAGGCCAGATTTCATATGAGCCCCCGGAATACAGCTTGAGCC 804
 QY 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 DB 805 GGTACAGGCCACCCACGCCAGTCGAAAGCTGTCAACCATCTCTTCCACAGTCCC 864
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 DB 865 AAAACTGAAAGACCGCTCTCTCAGTATGATCTATAGATTCATGAGACCAACAAAGTAGC 924
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeuGluLeu 320
 DB 925 CGCCTTCAATTCAGGCACTGGCCAGATCCAGCTTTGGAGATTCCTCCGAGACTCTCG 984
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphelysMet 340
 DB 985 TCGGACAGCTTCAACTCAGCTGATCACTGGAGAGGCCAACAAGGGGATTCAGATG 1044

QY 341 ThrAspProAspGluValAlaArgArgTrpGlyLlnArgLysSerLysProAsnMetAsn 360
 DB 1045 ACGGATCCCGCAGGAGTGGCCCGGCTGGGAGAGCGGAAGCAACCAACCATATGAC 1104
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrVal 380
 DB 1105 TACGATAGCTCACCGCCGCTCCGTACTACTATACAAAGAACATCATGACCAAGTGC 1164
 QY 381 HisProGlnSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 DB 1164 ----- 1164
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 DB 1165 CATGGGAAGCGCTACGGCTACAAAGTTGATTCACCGAGATGCCCGCGGCTCCAGCCC 1224
 QY 420 ----- 420
 DB 1225 CACCCCGGAGTCACTCTGTACAAGTACCCCTCAGACCTCCGTAACATGGCTCTCTAT 1284
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 DB 1285 CACGCCACCCACAGAGATGAATTTGTGGCGGCCACCTCCAGCCCTCCCGGTAC 1344
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysTyr 460
 DB 1345 TCTTCCAGATTTTGTGTGCCCCCAACCCATCTGGAATTCACCACTGGGGGTATATAC 1404
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 DB 1405 CCCACACTAGGCTCCCAACCAAGCATATGCTTCTCATCGGCGACTACTAC 1458

RESULT 2
 US-60-452-680-11396
 : Sequence 11396, Application US/60452680
 : GENERAL INFORMATION:
 : APPLICANT: CARGILU, Michele
 : APPLICANT: GROPE, Andrew
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 : TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 : FILE REFERENCE: CU001450
 : CURRENT APPLICATION NUMBER: US/60/452, 680
 : CURRENT FILING DATE: 2003-03-07
 : NUMBER OF SEQ ID NOS: 116213
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 11396
 : LENGTH: 1930
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-60-452-680-11396

Alignment Scores:
 Pred. No.: 5,34e-259 Length: 1930
 Score: 2324.50 Matches: 436
 Percent Similarity: 89.36% Conservative: 9
 Best Local Similarity: 87.55% Mismatches: 12
 Query Match: 89.82% Indels: 41
 DB: 18 Gaps: 3

US-09-902-772-4 (1-478) x US-60-452-680-11396 (1-1930)

QY 2 AlaSerThrIleuysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
 DB 25 GCACCTCATATCAAGGAGGAGGCTTATGAGTGTGAGTGAGGACCAAGTGTGTTGAGAGCT 84
 QY 22 AlATyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGluTyr 41
 DB 85 GCCTACGGAGACGCCACACCTGCTAGACAGATGACCGGCTCTCTCCAGCGACAT 144
 QY 42 GlyIlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
 DB 145 GGACAGACTTCCAAAGATGAGCCACGGCTCCCTCAGCAGATGGCTGTCAACCCCA 204

QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 Db 205 GCCAGGACACCATCAATAATGATGTAAACCTTAGCCAGGTGATGGCTCAAGGAACTCT 264
 QY 82 ProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnValGly 101
 Db 265 CCTGATCAATGCACTGGGCGCAAGGCGGGAAGATGTGGGCGACCCGACACCGCTGGG 324
 QY 102 MetAsnGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThAsn 121
 Db 325 ATGACATACGCGACCTCATGAGAGAGAGACATGCAACCCCAACATATGACCAAGAC 384
 QY 122 GluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArgGln 141
 Db 385 GAGCGCAGAGTTATTCGTGCCAGCAGATCCTACGCTATGAGATGACAGCAAGTCGGGCG 444
 QY 142 TripleGluTrrPalaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
 Db 445 TGGCTGGAGTGGGCGGTGAAGAATATGCGCTTCCAAATGCAACATCTGTATATTCAG 504
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
 Db 505 AACATCATGGAGGAAGAACTGTGCAGATGCAAGAGACAGATTCCAGAGGCTCACCCCC 564
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuArgGluThrProLeuPro 201
 Db 565 AGCTACAAACGCGACATCTTCTCTCATCTCCACTACCTCAGAGAGACTCTCTTCCA 624
 QY 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
 Db 625 CATTTGACTTCAGATGATGTGATTAAGCCTTACAAACCTCCACAGGTTAATGCATGCT 684
 QY 222 ArgAspThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241
 Db 685 AGAAGACAGAGGGGTGACGCTTTATTTTCCAAATCTTCAGATATCTCTGAAGCTACG 744
 QY 242 GlnArgIleThrThrArgProAspLeuProTyrGluGlnIleAlaArgAspSerAlaTrrThr 261
 Db 745 CAAGAATTTACAACTAGCCAGATTTACATATGAGCCCCCAGAGAAATCAGCTTGAGCC 804
 QY 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 Db 805 GGTCCACGCGCACCCACGCCCGCCAGTGCAGAAAGCTCTCAACCATCTCTCCACAGTCCC 864
 QY 281 LysThrGlnLysGlnArgProGlnLeuAspProTrrGlnIleLeuGlyProThrSerSer 300
 Db 865 AAAAAGTGAAGACAGGCTCTCTCACTAGTATCTTATCAGATCTTGGACCAACAAGTAGC 924
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGluLeuLeu 320
 Db 925 CGCCTTGCAGAAATCCAGGACAGTGGCCAGATCCAGCTKTGGCAGTTCCCTCGGAGCTCTG 984
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTrrPglLysThrAsnGlyLysPheLysMet 340
 Db 985 TCGGACAGCTCCAACTCAGCTGATCAGCTGGAAGCAACAAGGGGAGTTCAAGATG 1044
 QY 341 ThrAspProAspGluValAlaArgArgTrrPglLysLysSerLysProAsnMetAsn 360
 Db 1045 ACGGATCCGACAGAGGTGGCGCGCTGGGGAAGGGAAGCAAAACCAACATATTAAC 1104
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTrrTyrTrrAspLysAsnIleMetThrLysVal 380
 Db 1105 TAGCATTAAGCTCAAGCGCGCTCCGTTACTACTATACAAAGAACATCATGACCAAGGTC 1164
 QY 381 HisProProGluSerSerMetTrrLysTrrProSerAspLeuProTrrMetSerSerTyr 400
 Db 1164 ----- 1164
 QY 401 HisGlyLysArgTrrAlaTrrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 Db 1165 CATGGAGAGCGCTACGCTTCAAGATTCGACTTCCACGGGATTCGCCCGCTCCACGCC 1224
 QY 420 ----- 420

Db 1225 CACCCCGGAGTCATCTCTGTACAAGTACACCTCCAGACCTCCGTAACATGGCTCCAT 1284
 QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 Db 1285 CACGCCACCCACAGGAAGATGAACTTTGTGGCGGCCACCTCCAGCGCTCCCGTGACA 1344
 QY 441 SerSerSerPhePheAlaAlaProAsnProTrrTrrAsnSerProThrGlyLysTrr 460
 Db 1345 TCTTCCAGTTTCTTTTGTCTGCCCCAACCCTACTGGAATTCACCACTGGGGGATrrTATC 1404
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTrrTyr 478
 Db 1405 CCCAAGCACTGAGGCTCCCGACAGCAGATATGCTTCATCTGGGCGACTTACTAC 1458
 RESULT 3
 US-10-170-235-28519
 : Sequence 28519, Application US/10170235
 : GENERAL INFORMATION:
 : APPLICANT: VENTER, J. Craig
 : TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU
 : : FILE REFERENCE: CL001380
 : : CURRENT APPLICATION NUMBER: US/10/170,235
 : : CURRENT FILING DATE: 2003-03-17
 : : NUMBER OF SEQ. ID NOS: 42514
 : : SEQ. ID NO 28519
 : : LENGTH: 3294
 : : TYPE: DNA
 : : ORGANISM: HUMAN
 : : US-10-170-235-28519
 Alignment Scores:
 Pred. No.: 1,48e-243 Length: 3294
 Score: 2195.50 Matches: 416
 Percent Similarity: 84.97% Conservative: 8
 Best Local Similarity: 83.37% Mismatches: 10
 Query Match: 84.83% Indels: 65
 DB: 15 Gaps: 4
 US-09-902-772-4 (1-478) x US-10-170-235-28519 (1-3294)
 QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 94 ATGGCCAGCAGCTATTAAAGAAAGCTTATCACTGTGTGATGAGACCACTGCTTTGAG 153
 QY 21 CysAlaTrrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 Db 154 TGTGCTTACGGAAGCCACACCTGGCTTAAGACAGATGACCGGCTCCTCCACAGGAC 213
 QY 41 TrrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrrPleuSerGlnPro 60
 Db 214 TATGGACAGACTTCAAGATGAGCCACGCGTCCCTCAGAGGATTTGGCTGTCAACCC 273
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 274 CCAAGCCAGGCTCACATCAAAATGAAATGTAAACCTTAGCCAGTGAATGGCTCAAGGAC 333
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 Db 334 TCTCCTGATGAATGACAGTGTGGCCAAAGGCGGAAGATGTGGGACACCCAGACACCGTT 393
 QY 101 GlyMetAsnTrrGlySerTrrMetGluGluLysHisIleProProProAsnMetThrThr 120
 Db 394 GGGATGAATrACGCGACCTCATGAGAGAACACATGCGACCCCAACATGACACAGC 453
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
 Db 454 AACGAGGCGCAGATTTATTCGTGCCAGCAGATCCTACGCTATGAGATGACAGCAATGCGG 513
 QY 141 GlnTrrPleuGluTrrPalaValLysGluTrrGlyLeuProAspValAspIleLeuPhe 160
 Db 514 CAGTGGCTGAGTGGGCGGTGAAGAATATGCGCTTCCAGACGTCACATCTGTATTC 573

QY 161 GlnAsnIleaspGlyLysGluLeuCysLysMetThrItyAspAspPheGlnIArgLeuThr 180
 Db 574 CAGAAATCGATGGAGGAGAACTGTGCAGATGACCAAGAGCAGCTTCAGAGGCTCAC 633
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuAlaGluThrProLeu 200
 Db 634 CCCAGCTACAAAGCGCGACATCTTCTCCACATCTCCACTCCAGAGACATCTCTTT 693
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 Db 694 CCACATTGACTTCAGATGATGTGATTAAGCTTACAAACCTCTCCAGCGTTAATGCAT 753
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 Db 754 GCTAGAAACACA----- 765
 QY 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTyr 260
 Db 766 -----GATTTACCAATGAGCCCCCAGAGAGATCAGCCTGG 801
 QY 261 ThrSerHisSerHisProThr---GlnSerIysAlaThrGlnProSerSerThrVal 279
 Db 802 ACCGCTCAGCGCCACCCAGCCAGTCGAAAGCTGCTCAACCATCTCTTCCACAGCTG 861
 QY 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSer 299
 Db 862 CCCAAACTGAAAGCACAGCTCTCAGATTAGATCTTATCAGATTCTTGACCAACAAGT 921
 QY 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrGlnPheLeuLeuGlu 319
 Db 922 AGCGCCCTGCAAAATCCAGGACGAGTCCAGATCCAGCTTGGCAGTCTCTCGAGCTC 981
 QY 320 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGlnPheLys 339
 Db 982 CTGTCGAGCAGCTCCAACTCAGCTGATCATCTGGAGGACACCAAGGGAGTTTCAG 1041
 QY 340 MethTrpAspProAspGluValAlaArgArgTrpGlyGluArgLysSerLysProAsnMet 359
 Db 1042 ATGACGAGATCCCGAGGAGTGGCCCGCGCGGAGAGACGCAACCAACCAACATG 1101
 QY 360 AsnTrpAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetLys 379
 Db 1102 AACACATAGAGCTACACCGCGCTCGTACTACTGATGACAAAGAACATCATACCAAG 1161
 QY 380 ValHisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399
 Db 1162 GTC----- 1164
 QY 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
 Db 1165 ---CATGGAAAGCGCTACGCGCTACAGTGTGACTTCCACGCGGATGCGCCAGCGCTCAG 1221
 QY 420 Pro----- 420
 Db 1222 CCCCACCCCCCGGATATCTCTGTACAGTACCCCTCCAGACTCCCGTACAGTGGCTCC 1281
 QY 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
 Db 1282 TATACAGCCCAACAGAGAGATGAACTTGTGGCGGCCACCTCCAGCGCTCCCGCTG 1341
 QY 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyIle 459
 Db 1342 ACATCTTCCAGTATTTTGTGTCGCCCAACCACTACGGAATTCACCAACTGGGGGTATA 1401
 QY 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyLysTyrTyr 478
 Db 1402 TACCCCAACACTAGGCTCCCAACGCAATATGCTTCTCATCTGGGCACTTACTAC 1458
 RESULT 4
 US-60-452-680-11397
 ; Sequence 11397, Application US/60452680
 ; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele
 ; APPLICANT: GROPE, Andrew
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1001450
 ; CURRENT APPLICATION NUMBER: US/60/452,680
 ; NUMBER OF INVENTION: 2003-03-07
 ; NUMBER OF SEQ ID NOS: 116213
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 11397
 ; LENGTH: 3294
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-60-452-680-11397
 Alignment Scores:
 Pred. No.: 1,48e-243 Length: 3294
 Score: 2195.50 Matches: 416
 Percent Similarity: 84.97% Conservative: 8
 Best Local Similarity: 83.37% Mismatches: 10
 Query Match: 84.83% Indels: 65
 DB: 18 Gaps: 4
 US-09-902-772-4 (1-478) x US-60-452-680-11397 (1-3294)
 QY 1 MetaLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 94 ATGCCAGCAGCACTATTAGAAAGCCTTATCAGTTGTGATGAGGACCAAGCTGTTGAG 153
 QY 21 CysAlaIArgLysSerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 Db 154 TGTGCTACAGGAAGCCACACCTGGCTTAACACAGATGACCCGCTCTCTCCAGGAC 213
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 Db 214 TATGACAGACTTCCAGAGTGAAGGCCACGCTCCTCAGCAGATTTGGCTGTCAACCC 273
 QY 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 274 CCAGCCAGGCTCACACATCAAAATGGAATGTAACCTTACCGCAGGATGCTCAAGGAC 333
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 Db 334 TCTCTGATGATATGCACTGTGGCCAAAGGGGGAAGTGTGGGACACACACCGTT 393
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluCysHisIleProProProAsnMetThrThr 120
 Db 394 GGGATGAACCTACGGCAGCTACATGAGGAGAACACATGCCACCCCAACATGACCAGC 453
 QY 121 AsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg 140
 Db 454 AACGAGGCGAGAGTATCTGTCGACGAGATCTTACGCTATGAGTACAGACATGTGGCG 513
 QY 141 GlnTrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 Db 514 CAGTGGCTGAGTGGGGGGTGAAGAATATGAGGCTTCCAAATGTCAACATCTGTATTC 573
 QY 161 GlnAsnIleaspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 Db 574 CAGAAATCGATGGAGGAGAACTGTGCAGATGACCAAGAGCAGCTTCAGAGGCTCAC 633
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisIleuHisTyrLeuAlaGluThrProLeu 200
 Db 634 CCCAGCTACAAAGCGCGACATCTTCTCCACATCTCCACTCCAGAGACATCTCTTT 693
 QY 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 Db 694 CCACATTGACTTCAGATGATGTGATTAAGCTTACAAACCTCTCCAGCGTTAATGCAT 753
 QY 221 AlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 Db 754 GCTAGAAACACA----- 765

```

Oy 241 ThrGlnArgIleThrThrArgProAspLeuProTyrgInGlnAlaIaArgSerAlaTrp 260
Db 766 -----GATTACCATATGAGCCCCCAGAGATGACAGCTGG 801
Oy 261 ThrSerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerThrVal 279
Db 802 ACCGGTACAGGCGACCCCGCCAGTCGAAAGCTGTCACCATCTCTTCCACAGATG 861
Oy 280 ProLysThrGluAspGlnArgProGlnLeuAspProTyrgInIleLeuGlyProThrSer 299
Db 862 CCGAAACTGAAGACCGCTGCTCAGTACATCTTATGATTCCTTATGAGTTCGACCAACAGT 921
Oy 300 SerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrPglInPheLeuLeuGlnLeu 319
Db 922 AGCCGCTTGCAATCCAGCGAGTGGCCAGATCCAGCTTGGCAGTTCTCTCGAGCTC 981
Oy 320 LeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphelys 339
Db 982 CTGTGCGACAGCTCCACACTGCGATCAGTACCTGGGAAAGCACCACCGGGAGTTCAG 1041
Oy 340 MethrAspProAspGluValAlaIaArgTrpPglGlyLuphelysSerLysProAsnMet 359
Db 1042 ATGACGGATCCCGAGAGGTGGCCGGCGTGGGAGAGCGAAAGACCAACCAACATG 1101
Oy 360 AsnTyraAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLys 379
Db 1102 AACACAGTATAGCTCAGCGCGGCTCGTACTACATGACAAAGAACATATGACCAAG 1161
Oy 380 ValHisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSer 399
Db 1162 GTC----- 1164
Oy 400 TyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGln 419
Db 1165 ---CATGGAGAGCGCTACGCTACAGTTCACATTCACCGGAGATGCCAGGCCCTCAG 1221
Oy 420 Pro----- 420
Db 1222 CCCCACCCCGGAGTCACTCTGTACAAATGTAACCCCTCAGACCTCCGTACATGGGCTCC 1281
Oy 421 ---HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProVal 439
Db 1282 TATACAGCGCCACCAACAAATGATGATTTGTGGCGCCACCCCTCCAGCCCTCCGCTG 1341
Oy 440 ThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyLe 459
Db 1342 ACATCTTCCACTTTTTCCTGCTGCCCCAACCAACCATCTGGAATTCACCAACTGGGGGTATA 1401
Oy 460 TyrProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1402 TACCCCAACTAGGCTGCCACACGACCATATGCTCTGCTGCGACTTACTAC 1458

```

RESULT 5
PCT-US02-04915-41

```

: Sequence 41, Application PC/TUS0204915
: GENERAL INFORMATION:
: APPLICANT: Murray, Richard
: APPLICANT: Glynn, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Aziz, Natasha
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
: FILE REFERENCE: 018501-006200PC
: CURRENT APPLICATION NUMBER: PCT/US02/04915
: CURRENT FILING DATE: 2002-02-14
: PRIOR APPLICATION NUMBER: US 09/7784,356
: PRIOR FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: US 09/791,390
: PRIOR FILING DATE: 2001-02-22
: PRIOR APPLICATION NUMBER: US 60/285,475
: PRIOR FILING DATE: 2002-04-19
: PRIOR APPLICATION NUMBER: US 60/310,025

```

```

: PRIOR FILING DATE: 2001-08-03
: PRIOR APPLICATION NUMBER: US 60/350,666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/334,244
: PRIOR FILING DATE: 2001-11-29
: NUMBER OF SEQ ID NOS: 230
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 41
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US02-04915-41

Alignment Scores:
Pred. No.: 1.75e-242 Length: 1389
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: Gaps: 4

US-09-902-772-4 (1-478) x PCT-US02-04915-41 (1-1389)

Oy 2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCys 21
Db 25 GCAGCTCATATCAAGGAAGCCTTATCACTTGTGATGAGACAGCTCTTGTGAGTGT 84
Oy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlyTyr 41
Db 85 GCCTACGGAACGCCACACTGCTGTAAGACAGATGACCCGCTCTCTCCAGGAGTAT 144
Oy 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db 145 GGACAGACTTCCAGATGAGGCCACGCTCCCTCAGCAGAGTGGCTGTCTCAACCCCA 204
Oy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGGCGTACCAATGAATGTAACCTTACGCGAGTGAATGGCTCAAGAACTCT 264
Oy 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAsnValGly 101
Db 265 CCTGATGATGACATGTGGCCAAAGCGGGAAGTGGGGGACCCACGACACCGTGGG 324
Oy 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrAsn 121
Db 325 ATGAACTACGGCAGCTACATGAGAGAGAGACATGACCCCAACATGACCAAGAAC 384
Oy 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 385 GAGCGCAGAGTTATCGTCCAGACAGATCCCTACGCTATGAGATGACAGACATGCGGAG 444
Oy 142 TrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuPheGln 161
Db 445 TGCGTGGAGTGGGGGGAAGAAATATGCGCTTCCACAGCTCAACATCTTGTATTCCAG 504
Oy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCCATGGGAAGGAACTGTGCAAGATGACCAAGACACTTCCAGAGGCTACACCCC 564
Oy 182 SerTyraAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluTrpProLeuPro 201
Db 565 AGTACAAAGCGCGACATCTCTCTCAGATCTCCACACTGACAGAGACTCTCTTCCA 624
Oy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 625 CATTTGACTTCAGATGATGTTGATTAAGCCTTAAACCTCCACGCTTAATGCAAGCT 684
Oy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 685 AGAAACACA----- 693
Oy 242 GlnArgIleThrThrArgProAspLeuProTyrgInGlnAlaIaArgSerAlaTrpThr 261

```

```

Db 694 -----GATTACCATATGACGCCCCAGAGATCAGCCTGAC 732
Oy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 723 GGTACAGGCGCAACCCAGCCCAAGTCAAGAACTGCTCAACCATCTCTCCACATGCGCC 792
Oy 281 LysThrGluAspGlnAthrProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db 793 AAAACTGAAGACCGAGCTCTCATGATTAGATCTTATCATGATTCATTGACCAACAAGTAGC 852
Oy 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeuLeu 320
Db 853 GCGCTTGCATAATCCAGCGATGCGGATCCAGATCCAGCTTGGAGTCTCTGAGCTCTG 912
Oy 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphLysMet 340
Db 913 TCGGACAGCTCCAACTCCAGTCGATCCTGCGGAGGACCAACGGGAGTTCAAGATG 972
Oy 341 ThrAspProAspGluValAlaArgArgTrpGlyLysSerLysProAsnMetAsn 360
Db 973 ACGGATCCCGACGAGGTGGCGCGGCTGGGAGAGCGGAAGACCAACCAACATGAGC 1032
Oy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
Db 1033 TACGATTAGCTCAGCGCGGCTCCGTTACTACTACTATGACAGAACATCATGACCAAGTTC 1092
Oy 381 HisProGluSerSerMetLysTyrTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1092 ----- 1092
Oy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1093 CATGGAGAGCGCTACGCTACAGTTCGACTTCCACAGGAGATCGCCACAGCGCTCCAGGCC 1152
Oy 420 ----- 420
Db 1153 CACCCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTAT 1212
Oy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1213 CACGCCCAACCAAGAGATGACTTGTGGCGGCCCGCTCCAGCGCTCCCGTGACA 1272
Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyLysTyr 460
Db 1273 TCTTCCAGTTTCTTGTGCGCCCAACCATCATGGAATTCACCAACCTGGGGGTATATAC 1332
Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1333 CCCAACACTAGGCTCCCGACGACATATGCTTCTCATCTGGGCACTTACTAC 1386

RESULT 6
US-10-211-462-41
: Sequence 41, Application US/10211462
: GENERAL INFORMATION:
: APPLICANT: Murray, Richard
: APPLICANT: Glynn, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Aziz, Natasha
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
: TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
: FILE REFERENCE: 018501-006200US
: CURRENT FILING DATE: US/10/211,462
: PRIOR APPLICATION NUMBER: US 09/784,356
: PRIOR FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: US 09/791,390
: PRIOR FILING DATE: 2001-02-22
: PRIOR APPLICATION NUMBER: US 60/310,025
: PRIOR FILING DATE: 2001-08-03
: PRIOR APPLICATION NUMBER: US 60/334,244
: PRIOR FILING DATE: 2001-11-29
: NUMBER OF SEQ ID NOS: 230

```

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 41
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-211-462-41

Alignment Scores:
Pred. No.: 1,75e-242 Length: 1389
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 14 Gaps: 4

US-09-902-772-4 (1-478) x US-10-211-462-41 (1-1389)
Oy 2 AlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlnCys 21
Db 25 GCAGCTCATATCAAGAGGAGCCCTTATCAGTTGTGAGTGGACCAAGTGTGTTGAGTGT 84
Oy 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerSerGlyTyr 41
Db 85 GCCTACGGAGACCCACACCTGCTAAGACAGATGACCGGCTCTCTCTCCAGCGACTAT 144
Oy 42 GlyLlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
Db 145 GGACAGACTTCCAAAGATGAGCCCGCGGCTCCAGCAGATGCTGCTGTCTCAACCCCA 204
Oy 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db 205 GCCAGGGGTACACCAAAATGGAATGTATACCTTACCCAGGTGAATGCTCAAGGAACTCT 264
Oy 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101
Db 265 CTTGATGATGCACTGTGGCCAAAGCGGGAGAGTGTGGACCCAGACACCGTGTGG 324
Oy 102 MetAsnTyrGlySerTyrMetGluGlnLysHisIleProProProAsnMetThrThrAsn 121
Db 325 ATGAATCAGGCGACACTCATGAGAGAGACAGACATGCGACCCCAAAACATGACACGAC 384
Oy 122 GluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArgGln 141
Db 385 GAGCGAGAGATTATCGGCCAGCAGATCTTACGCTATGAGTACAGACCATGTGGCGCAG 444
Oy 142 TrpLeuGlnTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPheGln 161
Db 445 TGGCTGAGTGGGGGGAAGATATATGGCTTCCAGACGTCACATCTTGTATTTCCAG 504
Oy 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db 505 AACATCGATGGAGAGAACTGTGCAGATGACCAAGACAGCACTCCAGAGGCTCACCC 564
Oy 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
Db 565 AGCTTACAACCGCGCATCTTCTTCCATCTCCACTACCTCAGAGAGACTCTCTTCCA 624
Oy 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db 625 CATTTGACTTCAGATGATGTGTAAGAGCCTTACAAACCTTCACAGCGTTAATGCAATGCT 684
Oy 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
Db 685 AGAAGACCA----- 693
Oy 242 GlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTrpThr 261
Db 694 -----GATTACCATATGAGCCCCAGAGATCAGCCTGAC 732
Oy 262 SerHisSerHisProThr---GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db 733 GGTACAGGCGCAACCCAGCCCAAGTCAAGAACTGCTCAACCATCTCTCCACATGCGCC 792

```

QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleuGlyProThrSerSer 300
 |||
 Db 793 AAAACTGAAGACGACGGCTCCACTTACATCTTATCGATCTTGGACCAAGATGAC 852
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGlnLeuLeu 320
 |||
 Db 853 CGCCTTGCAATTCAGGAGCAGTGGCCAGATCCAGCTTTGGAGTTCCCTCCGAGCTCTG 912
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphLeuMet 340
 |||
 Db 913 TCGGACAGCTCCAACTCCAGTCGATCACCTGGGAGGACCAACGGGGAGTTCAAGATG 972
 QY 341 ThrAspProAspGluValAlaArgArgTyrPglGluArgGlySerLysProAsnMetAsn 360
 |||
 Db 973 ACGGATCCCGACGAGAGTGGCCGGCTGGGGAGACGGAGAGCAAAACCAACATGAC 1032
 QY 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 |||
 Db 1033 TACGATTAAGTCACACCGCGCCCTCCGTACTACTATGACAGAACATCATGACCAAGGTC 1092
 QY 381 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 |||
 Db 1092 ----- 1092
 QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 |||
 Db 1093 CATGGAGAGCGCTACGCTTACAGTTCGACTTCCACGGGATGCCCGCCCTCCAGCCC 1152
 QY 420 ----- 420
 Db 1153 CACCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCGCTACATGGGCTCTAT 1212
 QY 421 HisLysHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 |||
 Db 1213 CACGCCCAACCCAGAGATGAATCTGTGGCCCAACCTCCAGCCCTCCCGTGACA 1272
 QY 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrGlyGlyIleTyr 460
 |||
 Db 1273 TCTTCAGTTTCTTTGGTGGCCCAACCCATACATGGAATTCACCAACGGGGGTATATAC 1332
 QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 |||
 Db 1333 CCCAACACTAGGCTCCGCCACAGCATATGCTTCTCATCTGGGACACTTACTAC 1386
 RESULT 7
 US-10-087-192-1127
 : Sequence 1127, Application US/10087192
 : GENERAL INFORMATION:
 : APPLICANT: Morris, David W.
 : APPLICANT: Engelhard, Eric K.
 : TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 : TITLE OF INVENTION: CANCER
 : FILE REFERENCE: 52945200122
 : CURRENT APPLICATION NUMBER: US/10/087, 192
 : PRIOR FILING DATE: 2002-03-01
 : PRIOR APPLICATION NUMBER: US 09/747, 377
 : PRIOR FILING DATE: 2000-12-22
 : PRIOR APPLICATION NUMBER: US 09/798, 586
 : PRIOR FILING DATE: 2001-03-02
 : NUMBER OF SEQ ID NOS: 2059
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1127
 : LENGTH: 1389
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-087-192-1127

DB: 15 Gaps: 4
 US-09-902-772-4 (1-478) x US-10-087-192-1127 (1-1389)
 QY 2 AlasThrIleLysGluAlaLeuSerValValSerGluAspGlnSerIleuPheGluCys 21
 |||
 Db 25 GCAGCTCATATCAAGAGAAAGCTTATCACTGTGTGATGAGGACCAAGTCTGTTGAGAGCT 84
 QY 22 AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41
 |||
 Db 85 GCTTACGAGAACGCCACACCTGCTTAAGACAGATGACGCGGTCTCTCCAGGACATAT 144
 QY 42 GlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnProPro 61
 |||
 Db 145 GGACAGACTTCCAAAGATGAGCCACGCTCCCTCAGGAGATGGCTGTCTCAACCCCCCA 204
 QY 62 AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
 |||
 Db 205 GCCAGGGTCACCATCAAAATGGAATGTAACTTACCCAGGATGGAATGGCTCAAGAACTCT 264
 QY 82 ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValIly 101
 |||
 Db 265 CCTGATGATGACAGTGTGGCCAAAGCGGGAAGATGTGTGGCAGACCCAGACACCGTTGG 324
 QY 102 MetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrAsn 121
 |||
 Db 325 ATGAACTAGGACGCTACATGAGAGAGACATGCCACCCCAACATATACACGAAAC 384
 QY 122 GluArgArgValIleValProAlaAspProThrIleuTyrSerThrAspHisValArgGln 141
 |||
 Db 385 GAGCGAGAGATTATCGTCCAGACAGATCTCAGCTATGATGATACAGACATGTGGGAG 444
 QY 142 TrpLeuGluTrpAlaValLysGlyLysGlyLeuProAspValAspIleLeuLeuPheGln 161
 |||
 Db 445 TGGCTGGAGTGGGGGGTAAAGAAATATGCGCTTCACAGCGTCAACATCTTGTATTCAG 504
 QY 162 AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
 |||
 Db 505 AACATCGATGGGAAGAACTGTGCAAGATGACCAAGAGACGACTTCCAGAGGCTCACCCC 564
 QY 182 SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
 |||
 Db 565 AGTAAACAGCGGACATCTCTCTCATCATCTCCACATCTCAGAGACATCTCTTCCA 624
 QY 202 HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
 |||
 Db 625 CATTTGACTTCAAGATGATGTGATTAACCTTAAACCTCCACAGGCTTAATCATCTCT 684
 QY 222 ArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAlaThr 241
 |||
 Db 685 AGAAGACCA----- 693
 QY 242 GlnArgIleThrThrArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaTrpThr 261
 |||
 Db 694 -----GATTTACCATATGAGCCCCCAGAGAGATGAGCTCGAGACC 732
 QY 262 SerHisSerHisProThr--GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 |||
 Db 733 GGTACAGGCGCACCCACGCCCCAGTCGAAACCTCTCAACATCTCTCCACAGAGTGGCC 792
 QY 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleuGlyProThrSerSer 300
 |||
 Db 793 AAAACTGAAGACGCGCTCCACTTGAATCTTATCGATCTTGGACCAAGATGAC 852
 QY 301 ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuGlnLeuLeu 320
 |||
 Db 853 CGCCTTGCAATTCAGGAGCAGTGGCCAGATCCAGCTTTGGAGTTCCCTCCGAGCTCTG 912
 QY 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyLuphLeuMet 340
 |||
 Db 913 TCGGACAGCTCCAACTCCAGTCGATCACCTGGGAGGACCAACGGGGAGTTCAAGATG 972
 QY 341 ThrAspProAspGluValAlaArgArgTyrPglGluArgGlySerLysProAsnMetAsn 360

Alignment Scores:
 Pred. No.: 1,75e-242 Length: 1389
 Score: 2180.50 Matches: 413
 Percent Similarity: 84.74% Conservative: 9
 Best Local Similarity: 82.93% Mismatches: 11
 Query Match: 84.25% Indels: 65


```

Db      973  ACGGATCCCGACGAGTGGCCCGCGCTGGGAGAGCGGAGCAACCAACATGTAAC 1032
Qy      361  TyrAspIysIeuSerArgAlaLeuArgTyrTyrTyrAspIysAsnIleMetThrLysVal 380
Db      1033  TACGATAGCTCACCCCGCCCTCCGTTACTACTAGTACCAAGAACATCATGACCAAGGTC 1092
Qy      381  HisProGluSerSerMetTyrLysTyrProSerAspIeuProTyrMetSerSerTyr 400
Db      1092  ----- 1092
Qy      401  HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db      1093  CATGGAGACGCGCTACGGCTACAGATTGCTCCACGAGGATGCGCCAGCGCTCCAGGCC 1152
Qy      420  ----- 420
Db      1153  CACCCCGGAGTATCTCTGTACAGTACCCCTCAGACCTCCCGTACATGGGCTCCAT 1212
Qy      421  HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db      1213  CACGCCACCCACAGAGATGAATCTTGTGGCGCCGCCACCTCCAGCCCTCCCGTGACA 1272
Qy      441  SerSerSerPhePheAlaAlaProAsnProTyrTyrPheSerSerProThrGlyIleTyr 460
Db      1273  TCTTCAGGTTTTTTGTGCGCCCAACCCATCTGGAATTCACCACTGGGGGTATATAC 1332
Qy      461  ProAsnThrArgIeuProAlaAlaHisMetProSerHisLeuGlyThrTyr 478
Db      1333  CCCAACACTAGCGTCCCGCCACGACCATATGCTTCTCATCTGGGCACTTACTAC 1386

```

RESULT 8
US-10-126-052A-330
Sequence 330, Application US/10126052A
GENERAL INFORMATION:

APPLICANT: Mutray, Richard
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
FILE REFERENCE: 018501-001530US
CURRENT APPLICATION NUMBER: US/10/126, 052A
CURRENT FILING DATE: 2002-04-18
PRIOR APPLICATION NUMBER: US 60/284, 770
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: US 60/290, 492
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: US 60/339, 245
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: US 60/350, 666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/334, 370
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/372, 246
PRIOR FILING DATE: 2002-04-12
NUMBER OF SEQ ID NOS: 691
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 330
LENGTH: 3166
TYPE: DNA
ORGANISM: Homo sapiens
US-10-126-052A-330

Alignment Scores:

Pred. No.: 7,69e-242 Length: 3166
Score: 2180.50 Matches: 413
Percent Similarity: 84.74% Conservative: 9
Best Local Similarity: 82.93% Mismatches: 11
Query Match: 84.25% Indels: 65
DB: 14 Gaps: 4

US-09-902-772-4 (1-478) x US-10-126-052A-330 (1-3166)

```

Qy      2  AlaSerThrIleLysGluAlaLeuSerValValSerGlnAspGlnSerIlePheGluCys 21
Db      281  GCAGCTCATATCAAGGAAGCCTTATCATGTTGTGATGAGGACCAAGTGTGTTGAGAGT 340
Qy      22  AlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGluTyr 41
Db      341  GCCTACGGAACGCCACCTGGCTTAACAGAGATGACCGGTCCTCTCCAGGACTAT 400
Qy      42  GlnGlnHisSerLysMetSerProArgValProGlnGlnAspTyrPheLysGlnProPro 61
Db      401  GGACAGACTTCCAGATGATGAGCCCGCGCTCCCTACGAGATTTGGCTGTCTCAACCCCA 460
Qy      62  AlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSer 81
Db      461  GCCAGGGTACCATCAAAATGAAATGAACTTGAACCTTACCCAGGATGATGCTCAAGAACT 520
Qy      82  ProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGly 101
Db      521  CCTGATGAAATGCACTGTGGCCAAAGCGGGAAGATGGTGGACCCAGACCCGTTGG 580
Qy      102  MetAsnTyrGlySerTyrMetGluGluLysHisIleProProAsnMetThrThrAsn 121
Db      581  ATGAACTACGGCAGCTACATGAGAGAGACAGACATGCCACCCCAAAACATGACACCAAC 640
Qy      122  GluArgArgValIleValProAlaAspProThrIleuTyrPheThrAspHisValArgGln 141
Db      641  GAGCGCAGAGTTATCGTCCAGCAGATCTTACCTATGAGATGACAGACCATGTGCGGAG 700
Qy      142  TrpLeuGluTyrPheValValLysGluTyrGlyLysProAspValAspIleLeuPheGln 161
Db      701  TGGCTGAGTGGGGGGGGAAGAATATGCGCTTCCACAGCCCAACATCTGTATTTCCAG 760
Qy      162  AsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrPro 181
Db      761  AACATCATGAGGAAGAACTGTGCAACATGACCAAGACAGCACTCCAGAGAGCTACACCC 820
Qy      182  SerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuPro 201
Db      821  AGCTACACAGCGCGCATCTTCTTCACATCTCCACATCTCCAGAGAGATCTCTTCCA 880
Qy      202  HisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAla 221
Db      881  CATTTGACTTCAGTATGATGTGATTAAGCCCTTACAAACCTCTCCAGCGTTAAATGACGCT 940
Qy      222  ArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThr 241
Db      941  AGAAGACACA----- 949
Qy      242  GlnArgIleThrThrArgProAspLeuProTyrGlnAlaArgArgSerAlaTyrThr 261
Db      950  -----GATTTACCATATGACGCCCCCAAGAGATGACGCTGAGCC 988
Qy      262  SerHisSerHisProThr-----GlnSerLysAlaThrGlnProSerSerSerThrValPro 280
Db      989  GGTTCAGCGCCACCCACGCGCCCAAGTGAAGAGCTCTCAACATCTCTTCCACAGTGGCC 1048
Qy      281  LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
Db      1049  AAAACTGAAGACAGCGCTCTCATGATTGATCATCTTATGACATCTTGGACCAACAAGTAGC 1108
Qy      301  ArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuThrPheLeuLeuGlnLeu 320
Db      1109  CGCCTTGCAAAATCCAGGACAGTGGCCAGATCCACCTTTGGAGATTCCTCCGAGGCTCTG 1168
Qy      321  SerAspSerSerAsnSerAsnCysIleThrTyrGluGlyThrAsnGlyGluPheLysMet 340
Db      1169  TCGGACAGCTTCCAACTCCACCTGCTGAGGAAAGCCCAACGGGGAGTTCAAGATG 1228
Qy      341  ThrAspProAspGluValAlaArgArgTyrGlyLysArgLysSerLysProAsnMetAsn 360
Db      1229  ACGGATCCGACAGGAGTGGCGCGCTGGGAGAGCGAAGCAACCAACCAATGATGAC 1288
Qy      361  TyrAspLysIeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380

```



```

|||||
Db 1289 TACATAGCTACACCCGCGCTCCGCTACTACTATGACAAGAACATCATGCCAGTGC 1348
QY 381 HisProProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
Db 1348 -----
QY 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
Db 1349 CATGGGAAGCGCTACGGCTACAACTTGAGCTCCACGAGGATCGGCCAGGCCCTCCACCCC 1408
QY 420 -----
Db 1409 CACCCCGGAGTCATCTGTACAGTACCCCTCAGACCTCCCGTACATAGGCTCTCAT 1468
QY 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
Db 1469 CACGCCCCACCCAGAAAGATGAATCTTGCGGCCGACCCCTCCAGCCCTCCCGCTACACA 1528
QY 441 SerSerPhePheAlaAlaProAsnProTyrTyrPheSerProThrGlyIleTyr 460
Db 1529 TCTTCAGTTTGTGCTGCGCCCAACCATCTGGAATTCACCAACTGGGGGTATTTAC 1588
QY 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
Db 1589 CCCAACACTAGGCTGCCACACGACCATATGCTTCATCTGCGGCACTTACTAC 1642

RESULT 9
US-10-087-192-1124
: Sequence 1124, Application US/10087192
: GENERAL INFORMATION:
: APPLICANT: Morris, David W.
: APPLICANT: Engelhard, Eric K.
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: CANCER
: FILE REFERENCE: 529452000122
: CURRENT APPLICATION NUMBER: US/10/087,192
: CURRENT FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 09/7747,377
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 2059
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1124
: LENGTH: 1297
: TYPE: DNA
: ORGANISM: Mus musculus
US-10-087-192-1124

Alignment Scores:
Pred. No.: 3 07e-231 Length: 1297
Score: 2083.50 Matches: 394
Percent Similarity: 82.11% Conservative: 10
Best Local Similarity: 80.08% Mismatches: 7
Query Match: 80.51% Indels: 81
Gaps: 3

US-09-902-772-4 (1-478) x US-10-087-192-1124 (1-1297)
QY 7 GluAlaLeuSerValValSerGluAspGlnSerLeuPheGluCysAlaTyrGlySerPro 26
Db 2 GAGCGCTTGTCAGTGTGAGCGAGGACGACGACTATTGAGTGTGCTTACGGAACGCA 61
QY 27 HisLeuAlaLysThrGlnMetThrAlaSerSerSerSerGlnTyrGlyGlnThrSerLys 46
Db 62 CACCTGGCTAAGAGAGATGACCGCATCTTCGATGATGAGCTATGCGCAGACATCCAG 121
QY 47 MetSerProArgValProGlnGlnAspTyrLeuSerGlnProProAlaArgValThrIle 66
Db 122 ATGAGTCCAGAGAGTCCCTCAGCAGAGACTGCTGTCTAAGCCCGCAGGAGGTATCATC 181
QY 67 LysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspAspCysSer 86

```

```

|||||
Db 182 AAGATGAGTGCACACCTTAGTACAGTAAATGTTCCAGAACTCAGTCACTGACTGACGT 241
QY 87 ValAlaLysGlyGlyLysMetValSerSerSerSerAsnValGlyMetAsnTyrGlySer 106
Db 242 GTGAACAAGGTGGGAAGATGTGTGGCAGCCCGGATCTCTGTGGGATGATGACTAGCGCAGC 301
QY 107 TyrMetGluGlnLysHisIleProProProAsnMetThrThrAsnGlnLysArgValIle 126
Db 302 TACATGAGAGAGAGATGTGTCCGCCCTCCCAATTTGACCAAAATGAGCGGAGGTGATC 361
QY 127 ValProAlaAspProThrLeuTyrPheThrAspHisValArgGlnTyrPheGlnTyrAla 146
Db 362 GTCCCTCAGATCTCTACTGTGTGAGACAGACATGCTCCGACAGTGGCTGAGTGGGCGC 421
QY 147 ValLysGlnTyrGlyLeuProAspValAspIleLeuLeuPheGlnAsnIleAspGlyLys 166
Db 422 GTGAAGAATATGGCTCCCTCGATGTGAGTGGACGCTTACTATTTCAGAAATTCGATGGGAAG 481
QY 167 GluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186
Db 482 GAGCTGTGCAAGATGACAAAGATGACTTCCAGCGGCTCAGCCGAGCTACAAATGCCGAC 541
QY 187 IleLeuLeuSerHisLeuHisTyrLeuArgGlnThrProLeuProHisLeuThrSerAsp 206
Db 542 ATCTCTCTCAGATCTCCATCTCAGAGAGACTCCCTCCACATCTGACTTCCGAT 601
QY 207 AspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAlaArgAsnThrGlyGly 226
Db 602 GACCTTATTAAGCTTTACAAACTCTCCACGGTTAATGATGCGCAACACAGGGAGGT 661
QY 227 AlaThrPheIlePheProAsnThrSerValTyrProGlnAlaThrGlnArgIleThrThr 246
Db 662 GCACCTTTATTTTCCAAATACTTCACTATATCCGAGTACGCAAAAGATTAACAAT 721
QY 247 ArgProAspLeuProTyrGlnGlnAlaArgArgSerAlaThrThrSerHisSerHisPro 266
Db 722 AGGCCA-----
QY 267 ThrGlnSerLysAlaThrGlnProSerSerSerThrValProLysThrGluAspGlnArg 286
Db 727 -----
QY 287 ProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArgLeuAlaAspProGly 306
Db 728 -----GATCTTACAGATCTCTGGACCGACGACTGCCCTTGTCTAATCCAGGT 778
QY 307 SerGlyGlnIleGlnLeuTyrPheGlnPheLeuLeuGlnLeuLeuSerAspSerSerAsnSer 326
Db 779 AGTGGCCAGATCCAGCTGTGGCAGTCTGCTCGCAACTCGTGCAGACAGCTCCACACTCC 838
QY 327 AsnGlyIleThrTrpGlnGlyThrAsnGlyGlyLysMetThrAspProAspGlnVal 346
Db 839 AACGTGATCTACCTGGGAGGACACCAAGGGAGTTGAAATGACAGACCCGGAGAGAGTGG 898
QY 347 AlaArgArgTrpGlyGlnArgLysSerLysProAsnMetAsnTyrAspLysLeuSerArg 366
Db 899 GCTCGCGCTGGGGGAGAGAGAGAGCAAGCCCAACTGAACTATGCAACGCTAGCCGC 958
QY 367 AlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysValHisProProGluSerSer 386
Db 959 GCCCTCGCTACTACTACGACAAATAATCATGATGACCAAGTG----- 1000
QY 387 MetTyrLysTyrProSerAspLeuProTyrMetSerSerTyrHisGlyLysArgTyrAla 406
Db 1001 -----CACGGGAAGCGCTTACGCC 1018
QY 407 TyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro----- 420
Db 1019 TACAGTTTACTTCCAGGGATTTGCCAGGCTTGCAGGCCCACTCCCTGAGTGTCTCC 1078
QY 421 -----HisAlaHisProGlnLys 426

```

Db 1079 CTGTACAGTACCCCTCCGACCTGCATACATGGGCTCTTATCCAGCCGCCACCCAGAG 1138
 Oy 427 MetaspheValAlaProHisProAlaLeuProValThrSerSerPhePheAla 446
 Db 1139 ATGACCTTGTGTCTCCGACCCCTCCGCTCCAGACATCTTCAGATTCTTCT 1138
 Oy 447 AlaProAsnProTyrTrpAsnSerProThriglylyIleTyrProAsnThrArgLeuPro 466
 Db 1199 TCCCGAACCCATCTGGAATTCACGCACTGGGGGCACTCTACCCGAACTAGAGCTCCCA 1258
 Oy 467 AlaAlaHisMetProSerHisLeuGlyThrTyr 478
 Db 1259 GCCAGCCATATGCTCTCTACCTGGCACTACTAC 1294
 RESULT 10
 : Sequence 355, Application US/10144771
 : GENERAL INFORMATION:
 : APPLICANT: VENTER, J. Craig
 : TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
 : FILE REFERENCE: C1001321
 : CURRENT APPLICATION NUMBER: US/10/144,771
 : CURRENT FILING DATE: 2002-05-15
 : NUMBER OF SEQ ID NOS: 47235
 : SEQ ID NO 355
 : LENGTH: 2172
 : TYPE: DNA
 : ORGANISM: HUMAN
 : US-10-144-771-355
 Alignment Scores:
 Pred. No.: 3,77e-223 Length: 2172
 Score: 2017.50 Matches: 384
 Percent Similarity: 79.12% Conservative: 10
 Best Local Similarity: 77.11% Mismatches: 7
 Query Match: 14 Gaps: 97
 DB: 14 Gaps: 3
 US-09-902-772-4 (1-478) x US-10-144-771-355 (1-2172)
 Oy 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
 Db 107 ATGGCCAGCACTATTAAAGAGCCTTGTCACTTGTGACGAGGACAGCTACTATTGAG 166
 Oy 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 Db 167 TGTGCTTACGAGCAAGCCACACTGCTAGACAGATGACCCGATCTCTCCAGTGAC 226
 Oy 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGlnPro 60
 Db 227 TATGGCCAGACATCCAAAGATGAGTCCAGAGTCCCTGACGAGGACTGGCTGTCTCAAGCC 286
 Oy 61 ProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 Db 287 CCAGCCAGGCTCACATCAGATGAGTGCACACCTTACAGTGAATGTTCCAGAGAC 346
 Oy 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 Db 347 TCACCTGATGAGTGCAGTGTGAACAAGGTGGAGAGATGGTGGACCCGAGATACGTG 406
 Oy 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
 Db 407 GGGGTGGCTTACGCGACCTACATGAGAGAGAGATGTGCCCTCCCAATATATACACA 466
 Oy 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrSerThrAspHisValArg 140
 Db 467 AATGAGGACAGATGATGCTCCCTGCGATCTTACTTGTGAGACACAGACATGTCCGA 526
 Oy 141 GlnTrpLeuGluTrpAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
 Db 527 CAGTGGCTGAGTGGGGGAGAAAGATATGAGCTCTCTCATGTGGAGCGCTTACTATT 586
 Oy 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspIleGlnArgLeuThr 180

Db 587 CAGATATCGATGGAAGAGAGACTGTGCAAGATGCAAAAGATGACTTCCAGCGCTCAGC 646
 Oy 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
 Db 647 CCGAGCTTACATGCGGACATTTCTCTCCATCTCCACTACCTCAGAGACATCCCTT 706
 Oy 201 ProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis 220
 Db 707 CCACATCTGACTTCCGATGAGCTGTGATAGGCTTTACAAAACCTCCACGGTTAATCAT 766
 Oy 221 AlaArgAsnThrGlyGlyAlaThrPheIlePheProAsnThrSerValTyrProGluAla 240
 Db 767 GCCAGAAACACAGGGGTGACGCTTATTTTCCAAATACCTGATATATCCGAAGCT 826
 Oy 241 ThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaTrp 260
 Db 827 ACGCAAGAAATTAACACTAGC----- 847
 Oy 261 ThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPro 280
 Db 847 ----- 847
 Oy 281 LysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSer 300
 Db 847 ----- 847
 Oy 301 ArgLeuAlaAsnProGlySerGlyIleGlnLeuThrPglInPheLeuLeuLeuLeu 320
 Db 848 -----CCAGTATGAGGACAGATCCAGCTGTGGCGTCTGTCTGAGACTCTG 895
 Oy 321 SerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGluPheLysMet 340
 Db 896 TCACACAGCTCCCAACTCCCACTGATCAGCTGGGAGGACCAAGGGGAGTTCAAGATG 955
 Oy 341 ThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysProAsnMetAsn 360
 Db 956 ACAAGCCCGGACGAGGTGGCTGGGCTGGGGGAGAGAGAACCAAGCCCAACATCAAC 1015
 Oy 361 TyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMetThrLysVal 380
 Db 1016 TATACAAAGCTACAGCGCGCTCCGCTACTACTACAGCAAAAACATCATGACCAAGTG 1075
 Oy 381 HisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMetSerSerTyr 400
 Db 1075 ----- 1075
 Oy 401 HisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAlaLeuGlnPro 420
 Db 1076 CACGGGAAGGCTACAGCTTACAGTGTGACTTCCACGGGATGCCAGGCCCTGCAGCCC 1135
 Oy 420 ----- 420
 Db 1136 CACCTCTGAGTGTGCTCCGTACAACTACCCCTCCGACCTGCCATCATGAGGCTCTAT 1195
 Oy 421 HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeuProValThr 440
 Db 1196 CACGCCACCCCGAGAGATGAATTTGTCTTCCACCCCTCCGCTCCAGTACCA 1255
 Oy 441 SerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProThrArgLysIleTyr 460
 Db 1256 TCTTCCAGTTCTTGTGCTCCCGAACCCATCTGGAATTCACCGACTGGGGCACTCTAC 1315
 Oy 461 ProAsnThrArgLeuProAlaAlaHisMetProSerHisLeuGlyThrTyrTyr 478
 Db 1316 CCGAAGCTAGGCTCCAGCCAGCCATATGCTCTACCTGGGACACTACTAC 1369
 RESULT 11
 : PCT-US02-41414-1210
 : Sequence 1210, Application PC/TUS0241414
 : GENERAL INFORMATION:
 : APPLICANT: Morris, David
 : APPLICANT: Engelhard, Eric

```

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71249/RMS/DCF
; CURRENT APPLICATION NUMBER: PCT/US02/41414
; CURRENT FILING DATE: 2002-12-26
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1210
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US02-41414-1210

Alignment Scores:
Pred. No.: 3,82e-161 Length: 1359
Score: 1482.00 Matches: 291
Percent Similarity: 68.65% Conservative: 55
Best Local Similarity: 57.74% Mismatches: 80
Query Match: 57.26% Gaps: 78
DB: 2 Gaps: 9

US-09-902-772-4 (1-478) x PCT-US02-41414-1210 (1-1359)

Oy 1 MetLaserThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
Db 1 ATGGACGGGACTATTAGAGAGCTGTGGGTGGAGCGAGCAGCATCCCTCTTGAC 60
Oy 21 CysAlaTyrgly---SerProHisLeuAlaLysThrLumetThrAlaSerSerSer 39
Db 61 TCAGCGTACGAGAGGCGAGCCCATCTCCCAAGGCGCAGATGACGCTCGGGAGTCT 120
Oy 40 GluTyrglyGlnThrSerLysMetSerProAlaValProGlnGlnAspTyrLeuSerGln 59
Db 121 GACTACGGGCGAGCCCGACCATGATCAACCCCTCCACGACGAGAGTGTGATCAATCAG 180
Oy 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
Db 181 ---CCAGTGGAGGCTACAGTCAGCGGAGATAT-----GACCACTAATGATGATCCAGG 231
Oy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
Db 232 GAGTCTCGGTTGAGCTGAGCGGTAGCAATGACGAAAGCTGTGGGCGGAGCGAGTCC 291
Oy 100 ValGlyMetAsnTyrglySerTyrMetGluLysHis---LleProProAsnMet 118
Db 292 AACCCCATGACTACACAGCTATATGAGCAGAAAGATGCCCCCTCTCCACATG 351
Oy 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138
Db 352 ACCACGACGAGAGAGAGTATGTCGCCGACAGCCCACTGTGACAGAGAGAT 411
Oy 139 ValArgGlnTyrLeuGluTyrAlaValLysGluTyrGlyLeuProAspValAspPheLeu 158
Db 412 GTGAGGCAATGCTGAGTGGCCATAAAGATGACACTGTATGAGATGACACATCC 471
Oy 159 LeuPheGlnAsnIleAspGlyLysGluLeuLysMetThrLysAspAspPheGlnArg 178
Db 472 TTTTCCAGAACATGATGATGACAGAACTGTGTAAAGAAAGAGAGACTTCTCCGC 531
Oy 179 LeuThrProSerTyrAsnAlaAspIleLeuSerHisLeuAsnIleTyrLeuArgGlnThr 198
Db 532 GCCACACCCCTCTACACACGGAAGTGTGTGTTCACACCTCAGTTACTTACGGAAGT 591
Oy 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
Db 592 TCACGTG----- 597
Oy 219 MethSAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
Db 598 -----CTGGCCTATATATACAACTCCACACC 624

```

```

Oy 239 GluAlaThrGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSer 258
Db 625 GACCAATCTTCACGATTAGTGTCAAAAGACCTTCTTATGACTGATCAGAGAGAGA 684
Oy 259 AlaTyrThrSerHisSerHisProThrGlnSerIleValAlaThrGln---ProSerSer 277
Db 685 GCTTGGGCGCAATTAACATGATTTGCGCTCAACAAAGATCTCTCCCTTGAGAGGGCACA 744
Oy 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297
Db 745 ACGATCACTAAGATATACAGACCAAGCGCCAGCCAGATCCGTATCAATCATCTGGCCCG 804
Oy 298 ThrSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeu 317
Db 805 ACCAGCACTCCCTTAGCCACCACTCGAAGCGGCGAGATCCAGCTTGCAATCTCTCTG 864
Oy 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTyrGluGlyThrAsnGlyGlu 337
Db 865 GAGCTGCTCTCCGACGAGCGCAAGCCAGCTGTATCACTGAGAGGGAGACCAAGCGGAG 924
Oy 338 PheLysMetThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysPro 357
Db 925 TTCAAAATGACGAGCCCGATGAGGTGGCCAGCGCTGGGCGAGCCGGAAGAACGCC 984
Oy 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
Db 985 AACATGAAATTACGACAGCTAGCCGCGCTCTCTTATGATGATTAACATTTATG 1044
Oy 378 ThrLysValHisProGluSerSerMetTyrLysTyrProSerAspLeuProTyrMet 397
Db 1045 ACCAAAGTG----- 1053
Oy 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
Db 1054 -----CACGGCAAAAGATATGCTTAACAATTTGACTTCACAGGATGCCAGGCT 1104
Oy 418 LeuGlnPro----- 420
Db 1105 CTGACGCCACATCCGACCGAATCGTCCATGTACAAATCCCTTGACATCTCTCTACATG 1164
Oy 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
Db 1165 CCTTCATACATGCCACACAGAGATGAACCTTGTGCCCTCCCATCATCATCTCCATG 1224
Oy 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTyrPasnSerProThrGly 457
Db 1225 CCTGTCACTTCTCCAGCTCTTTGGAGCCGATCACAATCTGAGACCTCCGCCACGGGG 1284
Oy 458 GlyIleTyrProAsn-----ThrArgLeuProAlaIleHisMetProSerHisLeu 474
Db 1285 GGATCTACCCCAACCCCAAGCTCCCGCATCTTACACCCAGCTCTCTCACACTTA 1344
Oy 475 GlyThrTyrTyr 478
Db 1345 GGCAGCTACTAC 1356

RESULT 12
PCT-US02-41414-1209
; Sequence 1209, Application PC/TUS0241414
; GENERAL INFORMATION:
; APPLICANT: Engelhard, David
; APPLICANT: Morris, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71249/RMS/DCF
; CURRENT APPLICATION NUMBER: PCT/US02/41414
; CURRENT FILING DATE: 2002-12-26
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1613
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 1209
; LENGTH: 2957
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-41414-1209

Alignment Scores:
Pred. No.: 1,55e-160      Length: 2957
Score: 1482.00           Matches: 291
Percent Similarity: 68.65%      Conservative: 55
Best Local Similarity: 57.74%    Mismatches: 80
Query Match: 57.26%           Indels: 78
DB: 2                      Gaps: 9

US-09-902-772-4 (1-478) x PCT-US02-41414-1209 (1-2957)

Oy 1 MetLaseThrIleLysGluAlaLeuSerValAlaSerGluAspGlnSerLeuPheGlu 20
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 ATGACGGGACTATTAGAGAGGCTGTGCGTGGAGCGACGACCATCCCTCTTGAC 232

Oy 21 CysAlaTyrGly---SerProHisLeuAlaLysThrGluMetThrAlaSerSerSer 39
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 TCACGCGTACGAGGCGGCGACCATCTCCCAAGCGCCAGATGACTCCCTCGGAGTCT 292

Oy 40 GluTyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrLeuSerGln 59
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 GACTACGCGGCGACCCCAAGATCAACCCCTCCACACGACGAGAGTGCATCATCAG 352

Oy 60 ProProAlaArgValThrIleLysMetGluGlyAsnProAsnGlnValAsnGlySerArg 79
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 ---CCAGTGAAGGTCACGTCACGCGGAGTAT-----GACCACTGAATGATCCAGG 403

Oy 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 404 GAGCTCCGGTGAAGTGTAGCAAAATGCAAGCAAGTGTGGGGGAGGAGAGTCC 463

Oy 100 ValGlyMetAsnTyrGlySerTyrMetGlnGlyLysHis---LleProProAsnMet 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 464 AACCCCAAGACTACACAGCTATATGAGCAGAAAGATGCCCCCTCCCAACATG 523

Oy 119 ThrThrAsnGluArgValIleValProAlaAspProThrLeuTyrSerThrAspHis 138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 524 ACCACCAACGAGAGAGATATCTGCCCGACACCCACACTGTGGACACAGAGAT 583

Oy 139 ValArgGlnTyrLeuGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeu 158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 584 GTGAGCAATGGCTGAGTGGCCATTAAGAGTACAGCTGTATGAGATGACACATCC 643

Oy 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 644 TTTTTCAGAACATGATGCGACAGAACTGTATAAATGAACAAGGAGACTTCTCCCGC 703

Oy 179 LeuThrProSerTyrAsnAlaAspIleLeuSerHisLysTyrLeuArgGluThr 198
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 704 GCCACCAACCTTACACACAGAGAGTGTGTTCACACCTCACTTACCTGAGGAAGT 763

Oy 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 764 TCACTG----- 769

Oy 219 MethIleAlaArgAsnThrGlyAlaThrPheLlePheProAsnThrSerValTyrPro 238
    :|| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 770 -----CTGGCCATATATAACCTCCACACC 796

Oy 239 GluAlaThrGlnArgIleThrThrArgProAspLeuProArgGlnAlaArgArgSer 258
    ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 797 GACCAATCCCTACAGATGAGTGTCAAGAAACCCCTTATATGCTCAGTCAGAGAGA 856

Oy 259 AlaTyrThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 857 GCTTGGGGCATATACATGAATTCCTGGCCTACAAAAAGTCTCCCTTGGAGGGGACAA 916

Oy 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyPro 297

```

```

Db 917 ACGATCACTAGAAATACAGACCAAGCGCCCGACAGATCCGTATCAATCTGGCCCG 976
Oy 298 ThrSerArgLeuAlaAsnProGlySerLysGlnIleGlnLeuThrPheLeuLeu 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 977 ACCAGCACTCCCTAGCCCAACCCCTGGAAGCGGAGATCCAGCTGTGGCAATTCCTCTG 1036
Oy 318 GluLeuLeuSerAspSerSerAsnSerAsnGlyIleThrTrpGluGlyThrAsnGlyGlu 337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1037 GAGCTGCTCTCCGACAGCGCCACAGCCAGCTGTATCACTGTGAGGGGACCAAGGGGAG 1096
Oy 338 PheLysMetThrAspProAspGluValAlaArgArgTyrGlyGluArgLysSerLysPro 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1097 TTCAAAATGACGAGCGCCGATGAGTGGCCGCGGCGTGGGCGAGCGGAAAAAGCAAGCCC 1156
Oy 358 AsnMetAsnTyrAspLysLeuSerArgAlaLeuArgTyrTyrTyrAspLysAsnIleMet 377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1157 AACATGAATTCGACCAACCTAGCCGGGCTCCCTGTTATTCATGTATTAACATTTATG 1216
Oy 378 ThrLysValHisProProGlnSerSerMetLysTyrProSerAspLeuProTyrMet 397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1217 ACCAAAGTG----- 1225
Oy 398 SerSerTyrHisGlyLysArgTyrAlaTyrLysPheAspPheHisGlyIleAlaGlnAla 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1226 -----CAGGCAAAAGATATGCTTACAAATTTGACTTCCACAGGCATTCGCCAGGCT 1276
Oy 418 LeuGlnPro----- 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1277 CTGCAGCCACATCCGACCGAGCTGCCATGTACAAAGTACCTTGTACATCTCTTACATG 1336
Oy 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1337 CCTTCCTACCATGCGCCACAGAGAGTGAATTTGGCTCCCTCCCATTCATCTCCATG 1396
Oy 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTyrTrpAsnSerProTyrGly 457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1397 CCTGTCACTTCTCCACTTCTTGTGAGCGCCACAAATTAATCTGAGACTCCGCCACGGG 1456
Oy 458 GlyIleTyrProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1457 GGAATCTACCCCAACCCCAACAGTCCCGCCCATCTTAACACCCAGTCCCTTCACTTA 1516
Oy 475 GlyThrTyrTyr 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1517 GGCAGCTACTAC 1528

RESULT 13
US-10-342-887-710
; Sequence 710, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dal, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 710
; LENGTH: 2957
; TYPE: DNA

```

ORGANISM: Homo sapiens
US-10-342-887-710

Alignment Scores:

Pred. No.: 1.55e-160 Length: 2957
Score: 1482.00 Matches: 291
Percent Similarity: 68.65% Conservative: 55
Best Local Similarity: 57.74% Mismatches: 80
Query Match: 57.26% Indels: 78
DB: 14 Gaps: 9

US-09-902-772-4 (1-478) x US-10-342-887-710 (1-2957)

QY 1 MetAlaSerThrIleLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 173 ATGACGCGGACTATTAAAGAGCCCTGCGGTGTGAGCAGCAGCAGCTCCCTTTGAC 232
QY 21 CysAlaIArgIly---SerProHisLeuAlaLysThrIleMetThrAlaSerSerSer 39
DB 233 TCACCGTACGAGCGGCGGAGCCCATCTCCCAAGGCCACATGACTCGCTCGGAGGCT 292
QY 40 GluThrGlnIleThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGln 59
DB 293 GACTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 352
QY 60 ProProAlaArgValThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArg 79
DB 353 ---CCAGTACGAGGCGTCAACGTCACGCGGAGAT-----GACCAATGATGATCCAGG 403
QY 80 AsnSerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsn 99
DB 404 GAGCTCCGCGTGTGAGCTGAGCTGAGCAATGACAGCAAGCTGCGGCGGCGGAGTCC 463
QY 100 ValGlyMetAsnTrpLysSerTrpMetGluGluLysHis---IleProProAsnMet 118
DB 464 AACCCCATGACATCAACAGCTATATGAGAGAGATGAGCCCTCTCCCAACATG 523
QY 119 ThrThrAsnGluArgArgValIleValProAlaAspProThrLeuTrpSerThrAspHis 138
DB 524 ACCACACGAGAGAGAGATCATGCTCCCGCAGACCCACAGCTGTGACACAGACAT 583
QY 139 ValArgGlnTrpLeuGluTrpAlaValLysGluTrpLysLeuProAspValAspIleLeu 158
DB 584 GTGAGCGAATGCTGAGTGGCGCCATTAAGAGTACGCTTGATGAGATCGACACATCC 643
QY 159 LeuPheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArg 178
DB 644 TTTTCCAGAACATGATGCGCAGAGACTGTATAATGACACAGAGGAGACTTCTCCGC 703
QY 179 LeuThrProSerTrpAsnAlaAspIleLeuSerHisLeuHisTrpLeuArgGluThr 198
DB 704 GCCACACCTCTACACACAGAGAGTGTGTGACACCTCATCTACCTCAGGAGAGAGT 763
QY 199 ProLeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeu 218
DB 764 TCACTG----- 769
QY 219 MethAlaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro 238
DB 770 -----CTGCCCTATATTAACCTCCACACC 796
QY 239 GluAlaThrGlnArgIleThrArgProAspLeuProTrpGluGlnAlaArgArgSer 258
DB 797 GACCAATCTCAGCATGATGATGCAAGAGAACCTCTATATGCTCAGCAGAGAGAGA 856
QY 259 AlaTrpThrSerHisSerHisProThrGlnSerLysAlaThrGln---ProSerSerSer 277
DB 857 GCTTGGGCAATTAACATGATTCGCGCTCAACAAAGCTCCCTCGGAGGCGGCAAA 916
QY 278 ThrValProLysThrGluAspGlnArgProGlnLeuAspProTrpArgIleLeuGlyPro 297
DB 917 ACATTCAGTAAAGATACAGAGCAAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 976

QY 298 ThrSerSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu 317
DB 977 ACCAGCAGTCCCTTACAGCAACCTCGAAGCGGAGAGATCCAGCTGTGCAATTTCTCTG 1036
QY 318 GluLeuLeuSerAspSerSerAsnSerAsnCysIleThrTrpGluGlyThrAsnGlyGlu 337
DB 1037 GAGCTGCTTCCGACAGGCGCAGCAGCTGTATCACTGTGGAGGAGGAGGAGGAGGAG 1096
QY 338 PheLysMetThrAspProAspGluValAlaArgArgTrpGlyLysArgLysSerLysPro 357
DB 1097 TTCAAATGACGAGCCCGATGAGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1156
QY 358 AsnMetAsnTrpAspLysLeuSerArgAlaLeuArgTrpTrpTrpAspLysAsnIleMet 377
DB 1157 AACATGATATTCAGCAAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1216
QY 378 ThrLysValHisProProGluSerSerMetLysTrpSerProSerAspLeuProTyrMet 397
DB 1217 ACCAAAGTG----- 1225
QY 398 SerSerTrpHisGlyLysArgTrpAlaTrpLysPheAspPheHisGlyIleAlaGlnAla 417
DB 1226 -----CAGGCAAAAGATATGCTTACAAATTTGACTTCCAGGCAATTGGCCAGGCT 1276
QY 418 LeuGlnPro----- 420
DB 1277 CTCGAGCGCAATCCGACGAGTGTGCATACAAAGTAAACCTCTGACATCTCCATACATG 1336
QY 421 -----HisAlaHisProGlnLysMetAsnPheValAlaProHisProProAlaLeu 437
DB 1337 CCTTCTTACACATGCGCCACGACAGAGAGTGAATCTTCCCTCCCATTCATCTCCATG 1396
QY 438 ProValThrSerSerSerPhePheAlaAlaProAsnProTrpTrpAsnSerProThrGly 457
DB 1397 CCTGTCACTTCTCCAGCTTCTTGGAGCGCATCAACAATGAGTGAACCTCCCGCAGGG 1456
QY 458 GlyIleTrpProAsn-----ThrArgLeuProAlaAlaHisMetProSerHisLeu 474
DB 1457 GGAATTCACCCCAACCCCAACGATCGCCCGCATCTTAACACCCAGCGGCTTACACTTA 1516
QY 475 GlyThrTrpTyr 478
DB 1517 GGCAGCTACATAC 1528

RESULT 14
PCT-US02-41414-1207
Sequence 1207, Application PC/TUS0241414
GENERAL INFORMATION:
APPLICANT: Morris, David
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71249/RMS/DCF
CURRENT FILING DATE: 2002-12-26
PCT-US02-41414
PRIOR APPLICATION NUMBER: US 09/7747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
NUMBER OF SEQ ID NOS: 1613
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1207
LENGTH: 1359
TYPE: DNA
ORGANISM: Mus musculus
PCT-US02-41414-1207

Alignment Scores:
Pred. No.: 3.26e-160 Length: 1359
Score: 1474.00 Matches: 291
Percent Similarity: 68.25% Conservative: 53
Best Local Similarity: 57.74% Mismatches: 82
Query Match: 56.96% Indels: 78
DB: 2 Gaps: 9

OY		40	GluUtyrGlyGlnThrSerIysMetSerProargValProGlnGlnAspTyrLeuSergIn	59
Db		365	GACTACGGGAGCCCCAACAAATCATCAACCCTCCGCCACCGCAGCAGGAATGGATCAACAG	424
OY		60	ProProAlaArgValThrlleIysMetGlcysAsnProAsnGlnValAsnGlySerArg	79
Db		425	--CCAGTAGAAGTCATAATGTCAAGCGGGAGTAT-----GACCACATGAATTGATCCAGG	475
OY		80	AsnSerProAspPaspCysSerValAlaLysGlyGlyIysMetValSerSerSeraspSn	99
Db		476	GAGTCGTCCGGGACTGCAGTGTCAGCAAATTAACAACGTGTGGCGGAGCCCAACC	535
OY		100	ValGlyMetAsnTyrgLySerTyrMetGluGlnIyrsHs---lleProProRoasmet	118
Db		536	AACCCCATGAACATTAATAGCTACATGGATGAGAAGAACGGCCCCCTCTCTCCACATG	595
OY		119	ThrThrAsnGluArgArgValIlleValProAlaaspProthrLeutripSerThrAsphIs	138
Db		596	ACCACCAACGAGCAGGAGGATGATTGTGCTGCAGACCCACACTGTGGACACAGAGCAC	655
OY		139	ValArgGlnTrpLeuGlnTrpAlaValLysGluTyrgLyLeuProAspValAlspilleu	158
Db		656	GTTCGACAGTGCGCTGGAGTGGCGATAAAGATATCGGATGTATGATGACATTCACCTCC	715
OY		159	LeuphegInasnIleaspgLylysGluLeucysIysMetThrLysAspAspheGlnArg	178
Db		716	TTCITTCGCAACATATGGATGGCAGAGAATTTGTATAAATGACACMAGGAGACTTCTCCGA	775
OY		179	LeuthrProserTyrAsnAlaSpIleLeuLeuSerHisIleuHIStyrIleuaArgIluThr	198
Db		776	GCCACCTCCGCGCTCAACACACAGAGTGCTGTTCGCACCTCAGTATTCCTCAGGGAAAGT	835
OY		199	ProleuproHIsleuthrSerAspAspAlaSpIysAlaleuGlnAsnSerProargLeu	218
Db		836	TCACTG-----	841
OY		219	MethIsalaArgAsnThrGlyAlaThrPheIlePheProAsnThrSerValTyrPro	238
Db		842	-----CTGGCGCTTMAACNACNAACCTCCCATACA	868
OY		239	GluAlaThrGlnArgIleThrThrArgProAspleuProTyrGluGlnAlaAlaArgSer	258
Db		869	GACCAGTCCCTACGACGTGAATGTCAAGAMACACCTCTTATGACTGTGACGAGAGAGA	928
OY		259	AlaIrrThrSerHisSerHisProthcInberIysAla---ThcInProSerSerSer	277
Db		929	GCATGGACAATTAATATGAACTGTGGCTCTCAACAAAACCTCTTCCTTGGAGATCAACAG	988
OY		278	ThrValProLysThrGlnAspGlnArgProGlnLeuAspProTyrGlnIlleuendIyPro	297
Db		989	ACCATGGGCAAGAACACTGACGACGCGGCCACGCCACAGATCTTATCAGATCTCTGGGGCCA	1048
OY		298	ThrsSerArgLeuAlaAsnProGlySerGlyGlnIleGlnLeuTrpGlnPheLeuLeu	317
Db		1049	ACCAAGACGCCCTCAGCAAAACCTGGGAGTGGGAGATCCAGACTGTGGCAGTTCTCTCTG	1108
OY		318	GluIeuLeuSerAspSerSerAsnsSerAsncysIleThrTrpGluGlyThrAsnGlyLu	337
Db		1109	GAATCTACTGTCCGACAGCGCCACAGCGCTGTATCACCCTGGAGGGGACCAACGGGAG	1168
OY		338	PheIysMetThrAspProAspGluValAlaArgArgTrpGlyGlnArgIysSerLysPro	357
Db		1169	TTCAAAATGAGAGACCTCGATGAGTGGCCAGCGCTGGGAGAGCGGAAGACCAAGCACCC	1228
OY		358	AsnMelAsnTyraapLysLeuSerArgAlaLeuArgTyryrTyraSpLysAsnIlleMet	377
Db		1229	AACATGATATTTGACAGACTGACGCGGCGCTCCGATCTACTATGACAAATAACATTTATG	1288
OY		378	ThrlsYsaVlHsProProGlnSerSerMetIyryIyTyrProSerAspLeuProIyMet	397
Db		1289	ACCAAAAGG-----	1297

QY	398	serSerTYHISGLYAsrGTyAlaTYrLysPheAsrPheHISGLYlLeAlaInla	417
Db	1298	-----CATGGCAAAAGGATGCTCAACAAGTTGACTTTCATGGCATTGCCCAAGCC	1344
QY	418	LeuGlnPro-----	420
Db	1349	CTGCAGCCACATCCAAACAGAGACATCCATGTACAGATGCCCTGTGATATGCTTCATCATG	1408
QY	421	-----HISAlaHISProGlnLysMetAsnPhenValAlaProHISProProAlaLeu	437
Db	1409	CCTTCCTCATGCCATGCCCATCAACAGAAAGGTGAACCTTTGTCCGGTCCACCCATCCCTCATG	1468
QY	438	ProValThrSerSerSerPhePhenAlaAlaProAsnProTYrTTPAsnSerProthrgly	457
Db	1469	CCTGCACCTCTCTCCAGCTCTTTGGACACACATCAACATGTGAGACCTCCCCCACTCTCT	1528
QY	458	GLYlIeTYrProAsn-----ThrArgLeuProAlaAlaHISMetProSerHISLeu	474
Db	1539	GGGATCATATCCAAACCCCAAGTGTCCTCCCGCATCTCAACACCCAGCTGCCTTCACACATT	1588
QY	475	GLYThrTYrTYr 478	
Db	1589	GGCAGCTACTAC 1600	

Search completed: July 28, 2003, 10:00:10
Job time : 1115.35 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2003, 04:15:54 ; Search time 2008.73 Seconds

(without alignments) 3853.895 Million cell updates/sec

Title: US-09-902-772-4

Perfect score: 2588

Sequence: 1 MASTIKREALSVSEDSQSLFE.....IYPNTRLPAAHMPHSLGTY 478

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Command line parameters:

-OMODEL=frame_plus_p2n.model -DEV=xlh
-O/cgn2.1/USPto.spool/US09902772/funat_23072003_093657_14892/app_query.fasta.1.1294
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09902772_EGCGN.1.1.2619_6runat_23072003_093657_14892 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQDEXT -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=150 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_hum:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1153	44.6	1045	13	BM456833
2	1141.5	44.1	880	14	BQ233264
3	1121	43.3	634	13	BM489636
4	1111.5	42.9	1089	14	BQ212127
5	1085	41.9	915	12	BG390291
6	1053	40.7	715	12	BG388025
7	1010	39.0	629	10	AM948986
8	992.5	38.4	928	12	BG259459
9	984	38.2	701	9	AU136709
10	984	38.0	642	13	BT173036
11	984	38.0	668	10	BB660034
12	980	37.9	592	13	BM088063
13	968	37.4	827	9	AJ456498
14	948	36.6	675	10	BB639043
15	932	36.0	629	10	BB577887
16	882.5	34.1	1074	14	BQ953977
17	868.5	33.6	520	10	BB233460
18	845.5	32.7	949	13	BT558012
19	806	31.1	1015	9	AL554108
20	793	30.6	522	12	BF566705
21	769	29.7	516	12	BG691566
22	721	27.9	645	10	BB656030
23	715	27.6	666	12	BG390572
24	707	27.3	619	10	AM421310
25	701	27.1	671	13	BT588530
26	691	26.7	553	12	BG256864
27	686.5	26.5	1245	9	A1794504
28	666	25.7	1245	14	BQ212133
29	653.5	25.3	1068	14	BQ062565
30	648.5	25.1	872	14	BQ232274
31	643.5	24.9	740	9	AJ456506
32	639	24.7	894	14	BQ439561
33	619	23.9	815	9	AJ446143
34	616.5	23.8	756	9	AJ445352
35	609	23.5	787	10	BE559913
36	600.5	23.2	874	14	BQ215477
37	597.5	23.1	947	14	BQ642921
38	596.5	23.0	587	9	AL598619
39	588	22.7	651	10	BB442803
40	581.5	22.5	770	12	BG701574
41	580	22.4	788	9	AJ456869
42	573	22.1	907	10	BE561128
43	567.5	21.9	810	10	BE560516
44	558.5	21.6	972	12	BG257780
45	556	21.5	674	10	BB238163

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
BM456833	LOCUS	BM456833	1045 bp	mRNA	linear	EST 05-FEB-2002				
BM456833	LOCUS	AGENCOURT_6404041 NIH_MGC_92 Homo sapiens	CDNA	clone	IMAGE:5583753					
BM456833	LOCUS	5' mRNA sequence.								
BM456833	LOCUS	BM456833.1 GI:18505873								
BM456833	LOCUS	EST.								
BM456833	LOCUS	human.								
BM456833	LOCUS	Homo sapiens								
BM456833	LOCUS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
BM456833	LOCUS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
BM456833	LOCUS	NIH-MGC http://mgc.nci.nih.gov/.								
BM456833	LOCUS	National Institutes of Health, Mammalian Gene Collection (MGC)								

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12347 row: f column: 10
High quality sequence stop: 669.
Location/Qualifiers

FEATURES

source

1. 1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5583753"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 300 a 290 c 230 g 225 t
ORIGIN

Alignment Scores:

Score: 1153.00 Length: 1045
Percent Similarity: 76.078 Matches: 236
Best Local Similarity: 72.398 Mismatches: 24
Query Match: 44.55% Indels: 55
DB: 13 Gaps: 7

US-09-902-772-4 (1-478) x BM456833 (1-1045)

OY 7 GUAUAAUSeValValSerGUAUspGInSerLeuPheGluCysAlaTyrGlySerPro 26
DB 190 GAAGCCTTATCAGTGTGAGTGAAGGACGACGCTGTGTGAGTGTGCTACGAGAACGCA 249
OY 27 HistuAlaYstThrGluMetThrAlaSerSerSerSerGluTyrGlyGlnThrSerlys 46
DB 250 CACCTGGCTAAGACAGAGATGACCGCTCTCCACGACTGAGCACTGACACTCCAG 309
OY 47 MetSerProArgValProGlnGlnAspTyrLeuSerGlnProProAlaArgValThrIle 66
DB 310 ATGAGCCACGCGCCCTCAGCAGATGGCTGTCTCAACCCAGCCAGGAGTCAACATC 369
OY 67 LysMetGluCysAsnProAsnGlnValAsnGlySerArgAsnSerProAspPcysSer 86
DB 370 AAAATGAAATGATACCCCTAGCCAGGTGATGCTCAAGGAACCTCTCCATGAAATGACAGT 429
OY 87 ValAlaLysGlyGlyLysMetValSerSerSerAspAsnValGlyMetAsnTyrGlySer 106
DB 430 GTGGCCAAAGGGGGGAGAGATGGTGGCAGCCCAACACCGTTGGAGTGAATACGAGCAGC 489
OY 107 TyrMetGluGluTyrSHisIleProProAsnMetThrThrAsnGluArgValIle 126
DB 490 TACATGAGAGGAGAACATCCACCCCAACATGACACGAGAACGAGCAGAGATTATC 549
OY 127 ValProAlaAspProThrLeuTyrPheSerThrAspHisValArgIleTyrPheGluTyrPala 146
DB 550 GTGCACAGATCTTACGCTATGAGTACAGACCATGATGGCGAGTGGCTGAGTGGGCG 609
OY 147 ValLysGluTyrGlyLeuProAspValAlaSpIleLeuLeuPheGlnAsnIleAspGlyLys 166
DB 610 GTGAAAGAAATATGGCTTCCAGACGTCACATCTTGTATTCCAGAACATGATGGGAG 669
OY 167 GluLeuCysLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAlaAsp 186
DB 670 GAACGTGCAAGATGACCAAGACGACTTCCAGAGGCTCACCCCGAGCTACAGCCGAC 729

OY 187 IleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeuProHisLeuThrSerAsp 206
DB 730 ATCCCTTCTCAGCATCTTACATCACTAGAGAGACTCCCTTCACATTTGACCTCAGAT 789
OY 207 AspValAspLysAlaLeuGlnAsnSerProArgLeuMetHisAlaArgAsnThrGlyLys 226
DB 790 GATGTTGATTAAGCTTACAA-MACCTCCCGGTTAAAGCATGTGAAGAACCC----- 842
OY 227 AlaThrPheIlePheProAsnThrSerValTyrProGluAlaThrGlnArgIleThrThr 246
DB 842 ----- 842
OY 247 ArgProAspLeuProTyrGluGlnAlaArgArgSerAlaTyrThrSerHisSerHisPro 266
DB 843 -----GATTACCATATGAGACCCCGGAGATGACGCTGACGAGTCAACCC 896
OY 267 ThrGlnSerLysAlaThrGlnProSerSerSerThrVal-----Prolys 281
DB 897 -----CAAGCCCAAGTGGAAAGCTGTCAACATCTCCCTTCACAAAGGCCCAAA 947
OY 282 ThrGluAspGlnArgProGlnLeuAspProTyrGlnIleLeuGlyProThrSerSerArg 301
DB 948 ACTGAA-----AGACCCAGC-----GGCCCTCAGCT----- 974
OY 302 LeuAlaAsnProGlySerGlyGlnIleGlnLeuTyrGlnPheLeuLeuGlnLeuSer 321
DB 975 ---TAAATCCC-----TTAATCAATACCTTTGG 1001
OY 322 AspSerSerAsnSerAsn 327
DB 1002 GAACCCCAATATTAAC 1019

RESULT 2
BO233264
LOCUS
DEFINITION
BO233264 880 bp mRNA linear EST 02-MAY-2002
AGENCOURT 7565836 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6041526
5' mRNA sequence.

ACCESSION
BO233264
VERSION
BO233264.1 GI:20414664
KEYWORDS
EST.

SOURCE

human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 880)
NIH-MGC <http://imgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13279 row: h column: 07
High quality sequence stop: 608.
Location/Qualifiers

FEATURES

source

1. 880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6041526"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 250 a 231 c 207 g 189 t 3 others
 ALIGNMENT SCORES:
 Pred. No.: 4e-105 Length: 880
 Score: 1141.50 Matches: 221
 Percent Similarity: 95.02% Conservative: 8
 Best Local Similarity: 91.70% Mismatches: 8
 Query Match: 44.11% Indels: 4
 DB: 14 Gaps: 1

US-09-902-772-4 (1-478) x BQ233264 (1-880)

QY 1 MetLaserThrIleValLeuSerValValSerGluSerLeuPheGlu 20
 DB 74 ATGGCCGACTATTAAAGGAGCTTATCAGTTGTGATGAGGACCACTGTTTGTGAG 133
 QY 21 CysAlaTyrGlySerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
 DB 134 TGTGCTACGAGACGCCACCTGCTAGACAGATGACCGGCTCTCTCCACGAC 193
 QY 41 TyrGlyGlnThrSerLysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60
 DB 194 TATGACAGACTTCCAGATGAGCCACGCGCTCCCTCAGCAGATTGGCTGTCAACCC 253
 QY 61 ProAlaIleValIleThrIleLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 254 CCAGCCAGGCTCACCATCAAAATGATGATACCTTACCGCTGATGATGCTCAAGGAC 313
 QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
 DB 314 TCTCTGATGATGATGACGAGTGTGCGCAAGAGATGAGTGGCAGACGCCACGACGTT 373
 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProProAsnMetThrThr 120
 DB 374 GGGATGACACTACGCGACGATCATGAGAGAGACATGCCACCCCAATGATGACACG 433
 QY 121 AsnGluIleValIleValIleProAlaAspProThrLeuTyrPserThrAspHisValArg 140
 DB 434 AACGAGCGCAGATGATGTCGCCAGCATCTACGCTATGATGATGACGACCATGTGGG 493
 QY 141 GlnTyrPleuGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 DB 494 CAGTGGCTGAGTGGGCGGTGAAGAATATGGCTTCCAGACGTCACATTTTATTTC 553
 QY 161 GlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 554 CAGAACATCGATGGGAGAGAACTGTGCAAGATGACCAAGGACGACTTCCAGAGGCTCAC 613
 QY 181 ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProLeu 200
 DB 614 CCCAGCTACACGCGGACATCTCTCTCACATCTCCACTACCTCAAGAGACATCTCTT 673
 QY 201 ProHisLeuThrSerAspAspValAspLysAla-LeuGlnAsnSerProArg-LeuMetH 220
 DB 674 CCACATTGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
 QY 220 isAlaArgAsnThrGlyLysAlaThrPheIlePheProAsn---ThrSerValTyr 237
 DB 734 ATGGCTAGAAACACAGGAGGCGGTCTTTTATTTTCCCAATATACCTCAGTATAT 790

RESULT 3
 LOCUS BM489636 634 bp mRNA linear EST 07-FEB-2002
 DEFINITION pgm2n.pk011.124 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cdna clone pgm2n.pk011.124 5' similar to sp|Q90837|ERG.CHICK TRANSCRIPT. REGULATOR ERG p16S0754 transcription factor erg - chicken emb|CA54404.1| (X77159) ERG [Gallus gallus], mRNA sequence.
 ACCESSION BM489636
 VERSION BM489636.1 GI:18610567
 KEYWORDS EST.

SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.
 REFERENCE 1 (bases 1 to 634)
 AUTHORS Cogburn, L.A. and Monsonego-Ornan, E.
 TITLE ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IRAFS Animal Genome Project
 JOURNAL unpublished (2002)
 COMMENT Contact: Larry A. Cogburn
 University of Delaware
 Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES
 source
 1..634
 /organism="Gallus gallus"
 /strain="Commercial broiler and Ottawa Res. Centre strains 90 & 21"
 /db_xref="taxon:9031"
 /clone="pgm2n.pk011.124"
 /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"
 /tissue_type="Breast muscle, leg muscle and epiphyseal growth plate"
 /dev_stage="Breast, leg: Embryo(d19); post-hatch(d1,3,5,7,9,11 weeks); growth plate(d,7d,14d post-hatch)"
 /lab_host="E. coli EMDH10B"
 /note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

BASE COUNT 192 a 164 c 138 g 140 t
 ALIGNMENT SCORES:
 Pred. No.: 2.92e-103 Length: 634
 Score: 1121.00 Matches: 210
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 43.32% Indels: 0
 DB: 13 Gaps: 0

US-09-902-772-4 (1-478) x BM489636 (1-634)

QY 120 ThrAsnGluArgValIleValIleProAlaAspProThrLeuTyrPserThrAspHisVal 139
 DB 3 ACCAATGACGACAAAGATTATGTGCGACATCTCAGTATGAGACGACGACCATGTA 62
 QY 140 ArgGlnTyrPleuGlnTyrPalaValLysGluTyrGlyLeuProAspValAspIleLeuPhe 159
 DB 63 CGGAGTGGCTGGGAGGCGAGTGAAGATGATGCTCCAGAGCTGAGACATCTTGTG 122
 QY 160 PheGlnAsnIleAspGlyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeu 179
 DB 123 TTCAGAAACATTGATGGAAGAAGATTGTGTAATAATGACCAAGAATGACTTCCAGAGACT 182
 QY 180 ThrProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrPro 199
 DB 183 ACGCGGAGCTATACGAGATATCTCTCTCAGACCTACCTCAGAGAGACATCTCT 242
 QY 200 LeuProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMet 219
 DB 243 CTTCCACATTTGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
 QY 220 HisAlaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGlu 239
 DB 303 CATGCTAGAAACACAGGAGGCGCCTTTTATTTTCCCAATATACATCATGTTTACCGAGA 362

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10429 row: m column: 13
High quality sequence stop: 820.

FEATURES
SOURCE
Location/Qualifiers
1..915
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4524852"
/clone_lib="NIH-MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORE6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."
BASE COUNT 246 a 252 c 227 g 190 t
ORIGIN

Alignment Scores:
Pred. No.: 2,27e-99 Length: 915
Score: 1085.00 Matches: 238
Percent Similarity: 77.26% Conservative: 10
Best Local Similarity: 74.14% Mismatches: 26
Query Match: 41.92% Indels: 51
Gaps: 4

US-09-902-772-4 (1-478) x BG390291 (1-915)

QY 1 MetAlaSerThrIleuLysGluAlaLeuSerValValSerGluAspGlnSerLeuPheGlu 20
DB 59 ATGGCCACCACTATTAAAGAACCTTATCAGTTGTGAGTGAAGACCACTCTCTCCACGCGAC 118
QY 21 CysAlaIleuLysSerProHisLeuAlaLysThrGluMetThrAlaSerSerSerGlu 40
DB 119 TGTGGCTAGGAAACCCACACCTGGCTGAAGACAGATGACCGGCTCTCTCCACGCGAC 178
QY 41 TyrGlnIleuThrSerLysMetSerProArgValProGlnGlnAspThrLeuSerGlnPro 60
DB 179 TATGACAGACCTTCCAAAGATGAGCCACGCGCTCCAGAGATTGGCTGTCTCAACCC 238
QY 61 ProAlaArgValThrIleuLysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
DB 239 CCAGCCAGGTCACCATCAAAATGGAATGAACCTACCAAGTCAATGGCTCAAGGAC 298
QY 81 SerProAspAspCysSerValAlaLysGlyLysMetValSerSerSerAspAsnVal 100
DB 299 TCTCTGTATGATGATGAGTGTGGCCAAAGCGGGAAGATGTGGGCGCCACGACCGGTT 358
QY 101 GlyMetAsnTyrGlySerTyrMetGluGluLysHisIleProProPioAsnMetThrThr 120
DB 359 GGGATGACCTACGCGAGCTACATGAGAGAGACATGCCACCC-CCAAACATGACCGCG 417
QY 121 AsnGluArgValIleValProAla-AspProThrLeuThrSerThrAspHisValAr 140
DB 418 AACGAGCGCAGATTATGCTCCAGCAGATTCCTACCTATGAGATGAGATGAGATGTCG 477

QY 140 glnIleuPheGluIleuThrAlaValLysGluTyrGlyLeuProAspValAspIleLeuLeuPhe 160
DB 478 GCAGTGGCTGGAGTGGGGGTAAGAAATATGGCTTCCAGACGTCACATCTTGTATT 537
QY 160 eGlnAsnIleAspGlyLysGluLeuLysSerThrLysAspAspPheGlnArgLeuThr 180
DB 538 CCAGAACATCGATGAGGAGAGACTGTGCAGATGACCAAGAGCAGCTTCCAGAGCTCAC 597
QY 180 rProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuArgGluThrPro-L 200
DB 598 CCCCAGCTACAAATGGCGACAT-CTTCTCTCAATCTCCACTCCACAGAGACATCTTC 656
QY 200 euProHisLeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProAsnGluMet 220
DB 657 TTCCACATTGACTTCAATGATGATGTTGATTAAGCTTACAAACTCTCCAGCGTTAAATGC 716
QY 220 lAlaArgAsnThrGlyLysAlaThrPheIlePheProAsnThrSerValTyrProGluA 240
DB 717 ATGCTAAGAACCA----- 730
QY 240 lArhGlnArgIleThrThrArgProAspLeuProTyrGluGlnAlaArgSerAlaAr 260
DB 731 -----GATTACA-TATGAGCCCCCGCAGAGATCAAGGC 763
QY 260 rPThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerThrValP 280
DB 764 TGGACGGTCACGG-CACCCACAGCCCAAGTCACATCTCTTCCACAGTCCCAA 822
QY 280 rOlyThrGlu-----AspGlnArgProG 288
DB 823 TGAACACAGGTCGTAATCATGATCTTGGACAGTTCGCTGCATATCGGGGGCGGA 882
QY 288 lIleuAspProTyrGlnIleuGlyProThrSerArgLeuAlaAsnProGlySer 307
DB 883 CTTTGGATCCC-----GGGCGGGGAGC-----CCCGGCTCC 914

RESULT 6 715 bp mRNA linear EST 12-MAR-2001
LOCUS BG388025 602412867P1 NIH-MGC_92 Homo sapiens cDNA clone IMAGE:4521233 5',
DEFINITION mRNA sequence.
ACCESSION BG388025
VERSION BG388025.1 GI:13281471
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10420 row: f column: 18
High quality sequence stop: 633.

FEATURES
SOURCE
Location/Qualifiers
1..715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4521233"
/clone_lib="NIH-MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORE6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed."

OY	126	l1eva1ProAlaaspProThleuTripeSerhRspHsVAlArgGInTripleuGIntrp	145
Db	379	ATCGTGCCAGACAGATCTTACCGCTATGAGTACAGACATGTGCGCACTGGCTGGATGG	438
OY	146	AlaVallylGsluTgYlTgYleuProAspValAspIleuLeuPheGlnAsnIleAspGly	165
Db	439	CGGGTGAAGAATATGCGCTTCCAGAGCTCAAACTCTGTATTCCAGAAATCATGATGG	498
OY	166	LysGluLeuCyLysMetThrLysAspAspPheGlnArgLeuThrProSerTyrAsnAla	185
Db	499	AAGGAACGTGTCAGATGATGACCAAGACACACATCTCCAGAGAGCGTACACCCAGCTAACACGCC	558
OY	186	AspIleuLeuSerHisIleuHisTyrLeuArgGInuThrProLeuProHisIleuThrSer	205
Db	559	GACATCCCTCTCTCATATCTCATACTCACTCACTCAGAGAGACTCCTTCCACATTTGACTCA	617
OY	206	AspAspValAsp	209
Db	618	GATGATGGTGTAT	629
RESULT 8			
LOCUS	BG259459	928 bp	mRNA
DEFINITION	602378556p1 NIH_MGC_92 Homo sapiens CDNA clone IMAGE:4509541 5',		linear
ACCESSION	BG259459		
VERSION	BG259459.1	GI:12769275	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bts-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLM10389 row: O column: 14 High quality sequence stop: 717. Location/Qualifiers 1. 928 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4509541" /clone_1lb="NIH_MGC_92" /tissue_type="embryonal carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: testis; Vector: pCMV-SPORT6; Site: 1; NotI; Site-2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: This is a NIH_MGC Library."		
BASE COUNT	265 a	242 c	220 g 201 t
ORIGIN			
Alignment Scores:			
Pred. No.:	5.51e-90	Length:	928
Score:	992.50	Matches:	221
Percent Similarity:	83.63%	Conservative:	14
Best Local Similarity:	78.65%	Mismatches:	34
Query Match:	38.35%	Indels:	17
DB:	12	Gaps:	4

QY	1	MetaIaSerPrIIeYgSLuAlaLeuSerAlValSerGIuAspGInSerLeuPheGlu	20
Db	88	ATGGCCACCACTATTAAAGGAAGCCCTTACAGTGTGTGAGTGAGGACAGTGTGTG-TTTGG	146
QY	21	CysAlaTyrGIySerProHISLeuAlaLysThrGluMetThrAlaSerSerSerGIu	40
Db	147	TGTGGCTACGGAAAGCCCACTGGTGTAAAGACAGATGACCGGCTCTCTCCAGGAC	206
QY	41	TyrGIyGlnThrSerLysMetSerProArgValProGlnGlnAspTrpLeuSerGIuPro	60
Db	207	TATGGACAGACTTCCAGATGAGGCCACAGCGCTCCCTCAGCAGATGTGGCTGTCAACCC	266
QY	61	ProAlaIatgValIThrLLeuMetGIuCysAsnProAsnGlnValAsnGlySerArgAsn	80
Db	267	CCAGCCACAGTC-ACCATCAAAATGGAATGTAAACCTTG6CCAGGTGAATGGCTCAAGGAA	325
QY	81	SerProAspAspCysSerValAlaLysGlyGlyLysMetValSerSerSerAspAspVal	100
Db	326	TCTCTGTATGATATGACAGTGTGGCCAAAGCGGGAGAGATGGTGGGACACCAACCGTT	385
QY	101	GIyMetAsnTyrGIySerTyrMetGlnGluLysHisIleProProProAsnMetLThrThr	120
Db	386	GGGATGAAGTACGGCAGCTACATGAGAGAGACATGCGACCCCAACATGATGCCAGC	445
QY	121	AsnGIuATgATgValIleValProAlaAspProThrLeuTrpSerThrAspHisValArg	140
Db	446	AACGAGGCCAAGATTATCTGCCAGCAGATCTTCAGGTATGGACTATGACCAACATGTCCGG	505
QY	141	GlnTrpLeuGlnTTrpAlaValLysGluTyrGIyLeuProAspValAspIleLeuLysPhe	160
Db	506	CAGTGGCTGGAGTGGGCGGTACAGAAATATGGCTTCCAGACGTCACATCTTGTATTC	565
QY	161	GlnAsnIleAspGIyLysGluLeuCysLysMetThrLysAspAspPheGlnArgLeuThr	180
Db	566	CAGAACATCATGTGGGAAGAACTGTGCAGATGTACCAAGAGCAGACTTCCAGAGCTCAC	625
QY	181	ProSerTyrAsnAlaAspIleLeuLeuSerHisLeuHisTyrLeuATgGIuThrProLeu	200
Db	626	CCAG--TTCAAATGGCGACATCTTCTCTCACATCTCCACTCAAGAGAACTCTCTT	683
QY	201	ProHis-LeuThrSerAspAspValAspLysAlaLeuGlnAsnSerProArgLeuMetHis	220
Db	684	GCACATTTTGGCTTACATGATGTGTGATTAAAGCTTACAAAATCTCCAGG-TTAAATGC	742
QY	220	SalAArgAsnThrGIyLysAlaThrPheIlePheProAsnThrSerValTyrProGluAl	240
Db	743	TGCTAGA--ACCCAGGGGTGTGGTTTAA-TTTCAAAATATATTGTTT--TCTGAAGC	795
QY	240	aThrGlnArgTLeuThrArgProAsnLeuProTyrGlnGlnAlaArgysSerAlaTr	260
Db	796	TACCAACAGATTA-----CCACTTGGCAGTGCAGAAAACCCGGTGAATCGTGTG	843
QY	260	pThrSerHisSerHisProThrGlnSerLysAlaThrGlnProSerSerSerThrValPr	280
Db	844	GCAGAAC-----CAAGTCCGTGGATGGCGCAATCAAGAAAAACAAGATGCC	888
QY	280	o 280	
Db	889	C 889	
RESULT 9			
LOCUS	AUI36709	701 bp	mRNA linear EST 02-AUG-2002
DEFINITION	AUI36709 PLACE1 Homo sapiens cdna clone PLACE1004911 5', mRNA		
ACCESSION	AUI36709		
VERSION	AUI36709.1	GI:10997248	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Euteleia; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 701)		

AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
 Location/Qualifiers
 1..701

BASE COUNT 199 a 185 c 166 g 145 t 6 others
 ORIGIN
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE1004911"
 /clone_lib="PLACEL"
 /tissue_type="placenta"
 /note="Vector: pME18SF13"

Alignment Scores:

Pred. No.: 1.03e-89 Length: 701
 Score: 988.00 Matches: 189
 Percent Similarity: 94.20% Conservative: 6
 Best Local Similarity: 91.30% Mismatches: 10
 Query Match: 38.18% Indels: 2
 Gaps: 0

US-09-902-772-4 (1-478) x A0136709 (1-701)

OY 1 MetAlaserThrIleIysGluAlaLeuSerValSerGluAspGlnSerLeuPheGlu 20
 DB 85 ATGCCAGCAGCTATTAAAGAGCCCTTATCACTGTGAGTGAGCAGCAGCTGTTGAG 144
 OY 21 CysAlaTyrgIysSerProHisLeuAlaIleThrGluMetThrAlaSerSerSerGlu 40
 DB 145 TGTGCTTACGAGACGCCACCTGCTGACAGACAGATGACCGGCTCTCCACAGCAC 204
 OY 41 TyrgIysGlnThrSerIysMetSerProArgValProGlnGlnAspTyrPleuSerGlnPro 60
 DB 205 TATGACAGACTTCCAGATGAGCCACCGCTCCCTCAGCAGATGGCTGTCTCAACCC 264
 OY 61 ProAlaArgValThrIleIysMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80
 DB 265 CCACCCAGGCGCACATCAATGAATGTAACTGACCGCCAGGTGAATGGCTCAAGGAC 324
 OY 81 SerProAspAspCysSerValAlaIleIysGlyValIysMetValSerSerSerAspAsnVal 100
 DB 325 TCTCTGTATGATCACTGCTGCGCAAGCGGGAAGATGGTGGCAGCCAGACACCGTT 384
 OY 101 GlyMetAsnTyrgIysSerTyrMetGluGluIleHisIleProProProAsnMetThrThr 120
 DB 385 GGGATGAACCTACGCGCAGCTACATGAGAGAGACATGCCACCCCAACATATACACAG 444
 OY 121 AsnGluArgArgValIleValProAlaAspProThrLeuThrPserThrAspHisValArg 140
 DB 445 AACGAGCGCAGATTATCTGTCGCCAGATCTCAGCTATGAGTACAGACATGTCGCG 504
 OY 141 GlnThrPleuGluTyrPalaValIleIysGluTyrGlyLeuProAspValAspIleLeuPhe 160
 DB 505 CAGTGGCTGGAGTGGCGGTGAAGAAATAGCCCTTCAGACGTCACATCTTGTATTC 564
 OY 161 GlnAsnIleAspGlyIysGluLeuCysIysMetThrLysAspAspPheGlnArgLeuThr 180
 DB 565 CAGAACTCATGTGGGAANGAATCTGTCAAGATGACCAAGAGCAGCACTTTCAGANGCTCAA 624

OY 180 rProSerTyraAsnAlaIleLeuLeuSerHisLeuHisTyrLeuArgGluThrProL 200
 DB 625 CCCAGCTACAGACCGCAGCATCTTCTGTACATCTCAGTCACTCACTCAAGAGACTCTC 684
 OY 200 euproHisLeuThrSer 205
 DB 685 TTTCACATTTGNCNTCA 701

RESULT 10

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

EST.

ORGANISM

human.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)
 Other-ESTs: le01d08.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@leop.harvard.edu
 Laboratory was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center for information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@lm.wustl.edu)
 Seq primer: -40RP from Glbco
 High quality sequence stop: 473.
 Location/Qualifiers
 1..642

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5085758"
 /clone_lib="HR85 Islet"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1:
 Note: Site:2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@im.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692.

BASE COUNT

ORIGIN

178 a 171 c 163 g 129 t 1 others

Alignment Scores:

Pred. No.: 2.3e-89 Length: 642
 Score: 984.00 Matches: 183
 Percent Similarity: 97.46% Conservative: 9
 Best Local Similarity: 92.89% Mismatches: 5
 Query Match: 38.02% Indels: 0
 DB: 13 Gaps: 0

US-09-902-772-4 (1-478) x B1713036 (1-642)
<p> QY 1 MetalasertHrIleYsgLlAlaleuSerValYalSerGlualSpGlnSerleuPheGlu 20 Db 52 ATGGCCACACTATTAAGGAAGCCCTTACAGTTGTGACTGAGAGACCACTCGTTGTTGAG 111 QY 21 CysAlaTyrGlySerProHisleuNalalysrHrgImetHrAlaSerSerSerGlu 40 Db 112 TGTGGCTAAGGAAGCCACACCTGGCTTAAGACAGAGATGACCGGCTCTCCACGGCAC 171 QY 41 TyrGlyGlnThrSerIleMetSerProArgValProGlnGlnAspTyrIleuSerGlnPro 60 Db 172 TATGACACGACTTCCAAATATAGACCCACGCTCCCTCAGCAGATGTGGCTGTCTAACCC 231 QY 61 ProAlaArgValThrIleYsMetGluCysAsnProAsnGlnValAsnGlySerArgAsn 80 Db 232 CCAGCAGAGGTCACCATCAATAATGGAATGTAACCTTACCCAGAGTGAATGGCTCAGGAAC 291 QY 81 SerProAspAspCysSerValAlaIleSgIlyIlySmetValSerSerSerAspAsnVal 100 Db 292 TCTCCCTGTAATGATGACAGTGTGGCCAAAGCGGGAAGATGTGGCGCCACACCTT 351 QY 101 GlyMetAsnTyrGlySerTyrMetGluGluIlyHisIleProProAsnMetHrThr 120 Db 352 GGGATGACACTACGCGACGCTACATGTAGAGAGAAGCACATGCCACCCCAACATGACACG 411 QY 121 AsnGluArgArgValIleValProAlaAspProThrIleuTyrSerThrAspHisValArg 140 Db 412 AACGAGCGCAGAGTTATGTGTGCGACAGATCTCAGCTATGGATGAGACACCATGTGGG 471 QY 141 GlnTrrPleuGluTrrPalValIlySglutYrGlyIleuProAspValAspIleuIleuPhe 160 Db 472 CAGTGGCTGAGTGGCGGTGAAGAAATGGCTTCCAGACGTCACATCTTGTATTTC 531 QY 161 GlnAsnIleAspGlyIlySgluIleuCysIlySmetHrIlyAspAspPheGluArgIleuThr 180 Db 532 CAGAACATCGATGGGAGGAGGACTGTGCAGATGACACGACGACTTCCAGAGGCTCACCC 591 QY 181 ProSerTyrAsnAlaAspIleIleuIleuSerHisIleuHisTyrLeuArgIleu 197 Db 592 CCCAGCTATNCACGCGACATCTTCTCTCACATCTCCACTACCTCAGGA 642 </p>
<p> RESULT 11 BB660034 LOCUS BB660034 RIKEN full-length enriched, 13 days embryo lung Mus DEFINITION musculus cDNA D430027l04 5', mRNA sequence. ACCESSION BB660034 VERSION BB660034.1 GI:16493855 KEYWORDS EST. SOURCE house mouse. ORGANISM Mus musculus REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. AUTHORS 1 (bases 1 to 668) Arikawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hikimoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda, M., Koyama,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki, D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyu,T., Muramatsu,M. and Hayashizaki,Y. RIKEN Mouse ESTs (Arikawa,T., et al. 2001) Unpublished (2001) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shenocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/ </p>

Db	200	TATGGCCAGACATCCAGATGATGCCAGATCCCTCCAGCAGAGACTGGCTGCTCAACGCC	259
Qy	61	ProhlaarqValIThrIleIysMeTcIuCsAsnProAsnGlnValAsnGlySerArgAsn	80
Db	260	CCACCACGGGGCCACCATTCAAGATGGAGCGCAACCCCTACTCGATGAAGTGTCCAGGAAAC	319
Qy	81	SerProAspAspCysSerValAlaIalysGlyIlyIysMeTValSerSerSerAspAsnVal	100
Db	320	TCACCTATGATGTCAGCTGCTGACAAAGAGTGGGAAGATGGTGGGAGGCCGATCTGTG	379
Qy	101	GlyMeLAsnTyrgIySerTyMetcIuGlnIuLysHisIleProProProAsnMethrThr	120
Db	380	GGGATGAGCTCCGGCAGCTACATGAGAGAGAAGCATGTGCCGCCCTCCAAATATGACACCA	439
Qy	121	AsnGluArgValIleValProAlaAspProThrIeuTyrSerThrAspHisAlaArg	140
Db	440	AATAGAGCGCAAGATGATCTGCCCGCAATCTTCTGTGGAGCAGACCATTCGCA	499
Qy	141	GlnTyrPleuGlnTyrPalValIysGlnTyrgIlyLeuProAspValAspIleLeuDeuPhe	160
Db	500	CAGTGGCTCGAGTGGGGCGGTGAAGAATATGGCCCTNCICGATGTGGAGCTCTTACTATT	559
Qy	161	GlnAsnIleAspGlyIysGlnLeuCysIysMethrIyAspAspPheGlnArgLeuThr	180
Db	560	CAGATATTCGATGGGAAGACCTGTGCACAAATGACAAAGAGATGATTCACACGGCGCTCAG	619
Qy	181	ProSerTyrrAsnAlaAspIleLeuDeuSerHisIeuHisTyrrIeuArg	196
Db	620	CCGAGCTACATGCCAGACATTCCTCTCTCAATCTACACTACCTACGCA	667

RESULT	12.
LOCUS	BM088063
DEFINITION	592 bp mRNA linear EST 19-NOV-2001
ACCESSION	U01436 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.
VERSION	BM088063
KEYWORDS	BM088063.1 GI:16938691
SOURCE	EST.
COW.	

ORGANISM	REFERENCE
<i>Bos taurus</i>	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae	
Bovidae; Bovinae; Bos.	
1 (pages 1 to 592)	
Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T.,	

TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA
JOURNAL	Libraries and construction of a gene index for cattle
MEDLINE	Genome Res. 11 (4), 626-630 (2001)
COMMENT	21180013
	Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel.: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGAG
Plate: 139 row: D column: 21
Seq primer: ATTTAGGTGACACTATAG.

FEATURES

```
/organism="Bos taurus"  
/db_xref="taxon:9913"  
/clone_lib="MARC 2BOV"  
/tissue_type="pooled"  
/lab_host="DH10B"
```

BASE COUNT	ORIGIN
152 a	177 c 159 g 104 t

Alignment scores:	
Pred. No.:	5.17e-89
Score:	980.00
Percent Similarity:	96.89%
Best Local Similarity:	93.78%
Query Match:	37.87%
DB:	13
	Gaps:
	0
	Indels:
	0
	Mismatches:
	6
	Conservative:
	18
	Matches:
	59

US-09-902-772-4 (1-478) x BM088063 (1-592)

Oy	1	MetAlaSerThrIrrIleYsgIuaIaIeuserValIaSerGIuaSpGIserIeuPhcIu	20
Db	12	ATGGCCAGCAGCTTTTAAGGAAGCGTTATAGTCGTGAGGAGACCAAGTCACTGGTTTGA	71
Oy	21	CysAlaTyrClyserProHisIeuaIaIaYsThnGIuMetThIaISerSerSerGIu	40
Db	72	TGGCGCTACGGCAGCGCGCAGCTTGGCCCAAGAGGAAATACCGCGCTCTCCGCGGCAC	13
Oy	41	TyrClyIuInIhrSerIysMetSerProArGValProGIuInASPTrIeuSerGIuPro	60
Db	132	TATGGACAGACATCCAGATGAGCCCGCGCTCCCTCAGACAGACTGGCTGTCTAACCC	19

Oy	6	ProIaArvAlaThrIleYlsMetGluCysAsnProAsnGlnValAsmIylserTrgAsn	80
Dd	192	CCAGCCCGGGTCACCACTCAAGATGGAAATGAACCCTTAACCAAGGTACAACGGCTCAAGGAAAC	25
Oy	81	SerProAspApCyssSerValAlaIalysGlyLysMeValSerSerSerSpsanVal	100
Dd	252	TCCCTCATATAGTGCACCATGCCCCAANAAGTGGAAGAATGTGTGGCACGCCCGGACACCGTC	31
Oy	101	GlyMetAsnTyglYserTyrMetGluGlnIuLySHisIleProProProAsnmethrThr	120

Db	312	GGGATGAACTACAGCAGCTCATGAGGAGAAAGACATCCACCCCAAACTGACACC	372
Qy	121	ASNGLUATGATGATYALLILEVALPROLIAAAPPROTHTLEUTRPSERTHTASPHISVALRG	144
Db	372	AACGAGCGCCGCTCATCTCCCTGGAGATCCCAACCTATGAGTACACACCATATPTCCGG	433
Qy	141	GLNTPLEUNGUTRPAIALAVALLYSGLUTRYGYLEUPROASPVALASPILEULEUPHE	160
Db	432	CAGGGCGTGGATGGGCGAGTCAAAAGAAATACGGGGCTTCGGAGCTTGACATCTTATTATTC	499

QY	161	GLASPLTASPLLYYSGIUAECYSYLSMETTRHYSASASPHGLARDLTHR	181
QY			
DB	492	CAGATATTCACGGAGGAGCGCTGTCAAGATATGACCAAGATGACTTCAGAGGCTCAC	55
QY	181	PROSETRYTASNAALASPLLEULEUSERTHLEUHS	193
QY			
DB	552	CCGAGCTACACGCTACATCCCTCTGTGTGCACCTCCAC	590

[illegible]

Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
1 (bases 1 to 827)
Buerstedde, J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)

THIS PAGE BLANK (USPTO)